# INDUCTIVE SENSORS

## FOR INDUSTRIAL APPLICATIONS

SORIN FERICEAN

VISIT…

LANZAROTE
Caliente.com

# Inductive Sensors for Industrial Applications

For a listing of recent titles in the
Artech House *Microelectromechanical Systems Series*,
turn to the back of this book.

# Inductive Sensors for Industrial Applications

Sorin Fericean

ARTECH
HOUSE

BOSTON | LONDON
artechhouse.com

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the U.S. Library of Congress

British Library Cataloguing in Publication Data
A catalog record for this book is available from the British Library.

ISBN 13: 978-1-63081-255-3

**Cover design by John Gomes**

**© 2019 Artech House**
**685 Canton Street**
**Norwood, MA**

All rights reserved. Printed and bound in the United States of America. No part of this book may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission in writing from the publisher.

All terms mentioned in this book that are known to be trademarks or service marks have been appropriately capitalized. Artech House cannot attest to the accuracy of this information. Use of a term in this book should not be regarded as affecting the validity of any trademark or service mark.

10 9 8 7 6 5 4 3 2 1

*To all people involved, who*
*without a doubt supported me,*
*openly and repeatedly*

# Contents

Introduction xv

CHAPTER 1
Basics of Inductive Sensors, Definition, and Conventions 1

1.1 General Sensor Definition 1
  1.1.1 Sensor or Transducer 1
  1.1.2 IEEE Sensor Definition and Block Diagram 2
  1.1.3 Inductive Sensor Definition and Functional Block Schematic 3
1.2 Types of Inductive Sensors and Specific Classification Criteria 3
  1.2.1 Contact or Contactless Detection 3
  1.2.2 Position versus Displacement 4
  1.2.3 Absolute or Incremental Reading 4
  1.2.4 Linear or Angular Configuration 5
1.3 Main Features of Inductive Sensors: Definitions and Typical Values 5
  1.3.1 Supply Conditions and Limitations 5
  1.3.2 Sensing Range, Zero and Span, and Hysteresis 6
  1.3.3 Sensitivity and Nonlinearity, Linearity Error 7
  1.3.4 Accuracy, Resolution, and Repeatability: Three Precision Criterions 9
  1.3.5 Drift, Temperature Effects, and Temperature Ranges 10
  1.3.6 Dynamic Specification, Response Time and Cutoff Frequency, and Turn On and Turn Off Times 11
  1.3.7 Analog, Binary or Digital, and Voltage or Current Output Types 12
  References 16

CHAPTER 2
Inductive Proximity Sensors: Standards, EMC/EMI, Safety, Reliability, and Availability 19

2.1 Specific Product Standards and Requirements 20
  2.1.1 International Electrotechnical Commission Standard IEC 60947-5-2 20
  2.1.2 International Electrotechnical Commission Standard IEC 60947-5-7 32
2.2 Basic and Specific EMC/EMI Standards 35
  2.2.1 Resilience against Electrostatic Discharges 38
  2.2.2 Resilience against Radiated Electromagnetic Fields 38

2.2.3 Resilience against Fast Transients: EFT, Burst 39
2.2.4 Resilience against Impulse Voltage (Surge) 40
2.2.5 Resilience against Common Mode Conducted Disturbances 41
2.2.6 Magnetic Field Immunity Test 42
2.2.7 Immunity to Voltage Dips and Interruptions 42
2.2.8 Summary of the EMC Test Conditions for Inductive Proximity Sensors 42
2.3 Shock and Vibration Requirements 42
2.4 International Protection Classification 45
2.5 Intrinsic Safety, Product Safety Certification 46
2.6 Reliability and Availability 47
2.6.1 Mean Time Between Failures, Mean Time to Failure, and Failure Rate and Availability 47
2.6.2 Highly Accelerated Life Test 49
References 50

CHAPTER 3

Inductive Sensors: Definitions, Main Types, and Market Share 51
3.1 Overview of the Sensor Classification 51
3.2 Specific Embedding of Inductive Sensors 53
3.3 Main Types of Inductive Sensors 54
3.3.1 Magnetoelastic Systems 54
3.3.2 Electrodynamic Systems 55
3.3.3 Electromagnetic Systems with Closed Magnetic Loop 58
3.3.4 Electromagnetic Systems with Open Magnetic Loop 66
3.3.5 Variable Differential Transformers 72
3.3.6 Systems Based on the Eddy Currents Evaluation 74
3.3.7 Variable Transformers: Microsyn, Synchro, and Resolver 79
3.3.8 Final Considerations of Main Inductive Sensor Categories 87
3.4 Global Inductive Sensor Market: Size, Share, Growth, Trends, and Forecast 88
3.4.1 Global Sensor Market 88
3.4.2 Global P&D Sensor Market 88
3.4.3 Global IS Market 90
References 93

CHAPTER 4

Inductive Sensing Elements—Evaluation Methods 95
4.1 Analytical Methods of ISE 95
4.1.1 Inductors: Inductance, Impedance, Admittance, and Immittance 95
4.1.2 Quality Factor of an Inductor 102
4.1.3 Impedance and Q-Factor of a Resonant Circuit 103
4.2 Measuring Methods to Evaluate the ISE 108
4.2.1 Experimental Methods 109

4.2.2 Measuring Methods Suitable to be implemented in ISE 111
4.3 Modern Computer-Assisted Analysis and Synthesis of ISE 114
4.3.1 Fundamentals of the Computer-Aided Electromagnetic Field Simulation 114
4.3.2 Field Simulation Software Tools 119
4.3.3 Simulation with ANSYS Maxwell Tool 120
4.3.4 Flow Chart of a Maxwell Field-Simulation Project: Concrete Example 122
References 131

CHAPTER 5
Inductive Sensing Elements: Practical Implementations 133
5.1 Fundamental Inductors: Solenoid and Toroid 133
5.1.1 The Solenoid 133
5.1.2 The Toroid 135
5.2 Wire-Wound Coils with Air Cores 136
5.3 Wire-Wound Coils with Magnetic Cores 137
5.3.1 Inductance of Wire-Wound Coils with Magnetic Cores 138
5.3.2 Core Factor and Effective Core Parameters 139
5.3.3 Losses Caused by Cores: Total Formula of the Inductor Impedance at Full Length 141
5.4 Printed Flat Spiral Coils 146
5.5 Integrated Coils on Silicon Substrate 155
5.6 Active Inductors, Gyrators 158
References 161

CHAPTER 6
Magnetic Materials for Cores and Plungers of ISEs 163
6.1 Ferrites 163
6.1.1 Ferrites: Classification, Definitions, and Properties 164
6.1.2 Overview of MnZn Ferrite Specifications: Cross-Reference List of Available MnZn Ferrites 174
6.1.3 Ferrite Core Manufacturing Process: Technical Core Types 174
6.2 Permaloy and Mu-Metals 181
6.3 Soft Iron Alloys 185
References 186

CHAPTER 7
Evaluation Electronics of the Inductive Sensors 187
7.1 Generic Functional Diagram of the Inductive Sensor's Evaluation Electronics 187
7.2 Inductive Sensor with Discrete Evaluation Electronics 190
7.3 Inductive Sensor with Integrated Evaluation Electronics 193
7.3.1 Overview of Bipolar Integration Technology 194
7.3.2 Overview of Complementary Metal-Oxide-Semiconductor Processes: Benchmarking CMOS versus Bipolar Technology 205

7.3.3 Evaluation Electronics of Inductive Sensors with Integrated Circuits 207
7.4 ASIC Implementations in the Evaluation Electronics for Inductive Sensors 211
7.4.1 Single-ASIC Implementations: The Classical Device TCA505 211
7.4.2 Multi-ASIC Versions 218
7.4.3 Systems on Chip SOC 223
7.5 Software-Defined Sensors: Fantasy or the Inductive Sensor of Tomorrow? 231
References 234

CHAPTER 8
Excitation and Evaluation of the Inductive Sensing Element: Oscillators 237

8.1 Theory of Resonant LC Circuits: Series versus Parallel 237
8.1.1 Characteristics of the Series Resonant LC Circuit 238
8.1.2 Characteristics of the Parallel Resonant LC Circuit 239
8.2 General Theory of the Oscillator 242
8.2.1 Harmonic Oscillator 243
8.2.2 Linear Oscillator with Losses 248
8.2.3 Oscillators with Loss Cancellation by Positive Feedback Operation 250
8.3 Convenient Types of LC Oscillators for Inductive Sensors 257
8.3.1 Transistor-Based, Positive-Feedback Oscillators 258
8.3.2 Ring-Circuit Oscillators with LC Dipole 265
8.3.3 Differential Amplifier Oscillators with LC Dipole 276
8.3.4 Bridge-Network Oscillators 280
8.3.5 Oscillators with Pulsing DC Current Excitation 283
8.3.6 Negative-Resistance Oscillators 285
8.4 Function Generators 305
8.4.1 Relaxation Oscillators 306
8.4.2 Self-Oscillating Function Generators 307
8.4.3 Timer-Chip NE555 309
8.5 Digitally Synthesized Function Generators 312
References 313

CHAPTER 9
Inductive Sensors: Signal Processing and Conditioning 315

9.1 Signal Amplifiers 315
9.1.1 Operational Amplifiers: Definition and Applications 316
9.1.2 Operational Amplifiers: Frequency Response, Stability, and Compensation 320
9.2 Precision AC/DC Signal Converters 329
9.2.1 Precision Rectifiers 330
9.2.2 Peak Detectors 333
9.2.3 Synchronous Rectifiers 337
9.3 Sample-and-Hold Systems 341

9.4 Signal Linearization, Linearization Methods 343
9.4.1 Analog Hardware-Based Linearization 345
9.4.2 Software-Based Linearization 347
9.4.3 Logic Hardware Linearization 349
9.4.4 Hardware-Software Mixed Approaches 350
9.4.5 Artificial Neural Networks Approaches 352
9.5 Comparators, Window Discriminators 353
9.6 Regenerative Comparators (Schmitt Trigger) 355
9.7 Phase-Locked Loop Circuits 357
9.7.1 PLL Concepts 359
9.7.2 Analog PLL: Architecture and Operation 359
9.7.3 PLL Linear Analysis, Stability 361
9.7.4 Digital PLLs 362
9.8 Digital-to-Analog and Analog-to-Digital Convertors 365
9.8.1 Digital-to-Analog Converters 365
9.8.2 Analog-to-Digital Converters 368
References 371

CHAPTER 10
Inductive Sensors: Output Signal Providing 373
10.1 Output Stages for Analog Inductive Sensors 373
10.1.1 Voltage Output Stages 373
10.1.2 Current Telemetry: Current Output Stages 375
10.1.3 Ratiometric Voltage Outputs 380
10.2 Output Drivers for Digital Inductive Sensors 381
10.2.1 Switched Inductive Loads, Voltage Clamps 382
10.2.2 Output Drivers with Commercial Parts 384
10.2.3 Monolithic Integrated Output Drivers in ASICs 388
References 390

CHAPTER 11
Inductive Sensors: Power Supply and Sensor Protections 391
11.1 Power Supply Circuits 391
11.1.1 Series Voltage Regulators/References 392
11.1.2 Shunt Voltage Regulators 396
11.2 Standard and Supplimentary Sensor Protection Functions 399
11.2.1 Open Wire Protection 400
11.2.2 Reversed Polarity Protection 400
11.2.3 Protection against High-Energetic Pulses (Surge) 402
References 406

CHAPTER 12
Inductive Sensors: Adjustment and Calibration 407
12.1 Traditional Sensor Trimming Procedures with Commercial Components 408
12.1.1 Trimmable Resistors 408

- 12.1.2 Rejustor 410
- 12.1.3 Manual Mechanical Potentiometers 411
- 12.1.4 Digital Potentiometers 413

12.2 Specific Programmable Electronic Devices used to Calibrate Inductive Sensors 416
- 12.2.1 Established Methods for the Trimming of Inductive Sensors 417
- 12.2.2 ASIC and ASIC Sections for Trimming of Inductive Sensors. 422
- References 433

## CHAPTER 13

## Inductive Sensors: Temperature Compensation 435

13.1 Temperature Sensing Devices 435
- 13.1.1 Passive Temperature Probes 435
- 13.1.2 Active Temperature-Dependent Circuits 439
- 13.1.3 Active Temperature-Independent References: Bandgap References 444

13.2 Theoretical Considerations Regarding the Temperature Behavior of ISEs 446

13.3 Improvement of the Temperature Behavior by Passive Temperature Compensations 450

13.4 Active Analog Hardware-Based Temperature Compensation Methods of ISs 451

13.5 Active Digital Temperature Compensation Methods of ISs 456
- References 463

## CHAPTER 14

## Intelligent Inductive Sensors: Networking 465

14.1 Basics of the Data Communication Systems 465
- 14.1.1 Communication Network Topologies 465
- 14.1.2 Network Access Procedures 467
- 14.1.3 Industrial Fieldbuses: Definition and Features 469
- 14.1.4 ISO/OSI Network Reference Model 471

14.2 Requirement Description for the Sensor and Actuator Communication Level 474

14.3 Intelligent Sensors 475

14.4 RS-232, RS-422, and RS-484 Interfaces 477
- 14.4.1 RS-232 477
- 14.4.2 RS-422 and RS-485 478

14.5 Synchronous Serial Buses for Sensors 479
- 14.5.1 4-Wire Serial Peripheral Interface 480
- 14.5.2 2-Wire Interintegrated Interface 480

14.6 AS Interface 482
- 14.6.1 AS Interface at a Glance 482
- 14.6.2 AS Interface Slave Specification 484
- 14.6.3 ISO Reference Model of the AS Interface 485

14.7 IO-Link an Up-to-Date Sensor Communication System 489

14.7.1 IO-Link at a Glance 489
14.7.2 ISO Reference Model of the IO-Link Interface 490
14.7.3 IO-Link Communication Flow 495
14.7.4 IO-Link Hardware 496
14.8 Conclusions 498
References 498

Key to the Symbols for Currents and Voltages Used Throughout the Book 501
Key to the Symbols for Electronic Components 501

Acronyms and Abbreviations 503

Nomenclature of Electromagnetic Quantities 509

About the Author 511

Index 513

# Introduction

The answer to the question, Why write a book exclusively about inductive sensors? has to do with the more than 25 years of experience the author has spent working on the design and application of these sensors. According to professional statistics, the inductive sensor, with its numerous types and versions, is very popular in the sensor world; it is produced by a vast number of small-, medium-, and large-sized companies, so that the total volume of inductive sensors sold globally places them in the top ten in the sensor market. The reasons for this are due to specific features such as a low cost for low, medium, and high performance, a medium sensing range with precision, high practicality reflected in the long service life, and good stability.

At the same time, there is no known technical or scientific book which covers all the facets of this classical device, from the operation fundamentals to a description of the base embodiments and types, including their architecture, functionality, and construction.

In this book, Chapters 1–3 give a general view of inductive sensors: topics such as the basics, definitions and conventions, features and standards, as well as classes and classifications are discussed. The goals are to present several definitions and conventions available in different technical publications, and to establish a lexicon based on the international standards which are valid worldwide and will systematically be used in this book. Chapter 3 dives into the world of inductive sensors and describes the most significant families.

Chapters 4–6 deal with the first key subsystem of an inductive sensor, the inductive sensing element. These chapters explain how to evaluate the primary information provided by these units, as well as how they are made (implementations, materials, etc.).

Chapter 7 is pivotal to the book, as it describes the evolution of the second key subsystem of an inductive sensor, that is, the evaluation electronics. The generic structure illustrated in Figure 7.1 shall serve as a compass which will guide the reader through the subsequent chapters.

Chapters 8–13 deal with operation methods and several circuits for the realization of the significant functionalities of all types of inductive sensors.

Finally, Chapter 14 focuses on sensor specific digital interfaces.

This book has some specific characteristics:

- The book addresses professionals—engineers and technicians—and also students with a major in electronics, as well as anyone who requires a solid basic knowledge of inductive sensors;

- For this reason, every chapter debuts with classic, traditional solutions and gradually moves on to state-of-the-art analog and digital realizations (large-scale, integrated, systems-on-chip, software-defined sensors (SDSs), digital signal synthesis, coils on silicon, active inductors, artificial neural network (ANN) approaches, etc.), which are implemented in various types of inductive sensors;
- The book systematically uses three conventional but also modern analysis methods: (a) the analytic computation (as far as applicable), (b) popular graphical methods (phasor diagrams, phase planes, Smith-charts, signal path, etc.), and (c) computer assisted tools, such as the ANSYS Maxwell electromagnetic field simulation software, and the most popular Spice simulator for electronic circuits. These tools offer a deeper view into the phenomenology of the analyzed devices (see Chapters 3 and 4);
- For traditional solutions, the chapters give overviews in tables with computation formulas (including empirical expressions) found in various sources (e.g., Table 5.1);
- Concrete examples for actual, available commercial parts are always provided; and
- Numerical examples to help the reader consolidate the theoretical knowledge acquired. Example 13.1 (Chapter 13) is to be specifically noted: It is a corollary design exercise which makes use of knowledge gained in the previous chapters.

In view of these features, it is easy to understand the intention of this book: to provide a comprehensive survey of all relevant inductive sensor classes for industrial applications in one single volume.

In other words

*All information in one book and one book for all readers.*

Finally, for a deeper insight into the applications of the inductive sensors, the author invites the reader to study *Sensoren in Wissenschaft und Technik,* Second Edition, E. Hering, and G. Schonfelder (eds.), Wiesbaden, Germany: Vieweg+Teubner Verlag/Springer Fachmedien GmbH, 2018. Good presentations of the popular commercial sensors are available here.

The author would like to thank Balluff GmbH in Germany, particularly Mr. Rolf Hermle, Mr. Ernst Gass, and Mr. Michael Unger, for providing excellent working conditions and permanent support throughout the long cooperation time. Special thanks to all assistant staff members of the Sensors Development Department–first and foremost Mr. Albert Dorneich—but also to external project cooperation partners for the permanent readiness for experiments and new findings.

Last but not least, the author would like to express his gratitude to all private persons who support him, morally or de facto.

CHAPTER 1

# Basics of Inductive Sensors, Definition, and Conventions

The purpose of this introductory chapter is to summarize and to classify the large number of specific terms and definitions relative to the inductive sensor (IS). The sensors for industrial applications are specified in national and/or international product and/or generic standards. These important documents cover, in part, the specific terminology, but are—if one is honest—not easy to read.

The goal of this chapter is to combine the standard descriptions with additional terms used in technical publications, and to define all in a uniform manner, in order to provide a complete, exhaustive technical vocabulary of sensors in general, and inductive sensors in particular.

## 1.1 General Sensor Definition

There are many terms that are often synonymously used for sensor, including transducer, detector, and so forth. A widespread and frequent source of confusion is the definition of a sensor and its functionality. What is a sensor? According to *The Electrical Engineering Handbook* [1]: "Defining the term sensor is not an easy task." Unfortunately, the confusion between sensors and transducers goes so deep there are technical events called either sensor- or transducer-conferences.

### 1.1.1 Sensor or Transducer

In order to clarify this major ambiguity, we refer to reference work [1]. According to this publication, the most widely used sensor definition is based on the definition that the Instrument Society of America has applied to electrical transducers, that is, "a device which provides a usable output in response to a specified measurand." The transducer converts energy from one to another. More exactly, it converts a signal from one physical form to a corresponding signal having a different physical form [2]. The definition refers to six different kinds of signals, namely mechanical, thermal, magnetic, electric, chemical, and radiation (including optical). A typical example is the classic wall thermometer, which converts temperature into mechanical deviation of a column of mercury.

The delimitation of the output to an electronic signal leads to the term sensor, from the Latin word sensus, meaning to sense. A sensor is generally defined as a device that provides a usable electronic output in response to a specific input quantity, called a measurand.

Obviously, according to these definitions, a transducer can sometimes be a sensor, and vice-versa. In this book, we will concern ourselves only with sensors.

### 1.1.2 IEEE Sensor Definition and Block Diagram

In the previous IEEE sensor definition, the measurand can be a physical, chemical, or biological property, which is evaluated. The conversion of the arbitrary measurand into a sensor electronic output signal employs one or more transduction mechanisms. The classic representation is shown in Figure 1.1.

Independent of the measurand kind, the output of the primary transduction mechanism, called the primary signal, is now a physical information, which can be evaluated by the second transduction mechanism and finally converted into an electrical signal. This general sensor definition is accepted worldwide.

Taking into consideration the final destination of a sensor, this definition can be expressed in more practical terms. The term sensor is used to mean a technical device, which converts the measurand into an unambiguous electronic sensor output signal. This signal is generally a voltage or a current and may be analog, binary, or digital encoded.

Finally, it is reasonable to conclude this theoretical consideration with an enumeration of the major advantages provided by sensors [2]:

1. Sensors can be designed for any nonelectric measurand—a very significant aspect for inductive sensors, whose inputs are geometrical or material characteristics, in general.
2. Energy does not need to be drained from the measurand. The sensor requires a power supply from the application (Figure 1.1) and—energetically speaking—modulates that energy.
3. There is a wide range of modern, low-power consumption integrated circuits (commercial or manufacturer-proprietary) available that are able to perform high performance transduction mechanisms.
4. Many options exist for information display or recording for electronic output signals.

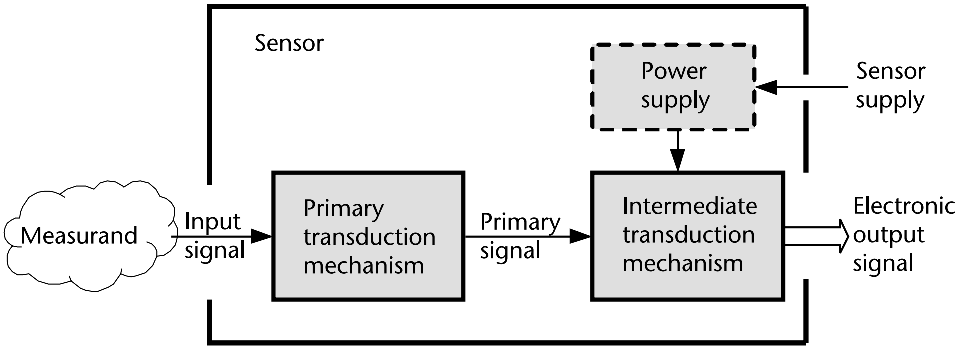


**Figure 1.1** General sensor definition according to [1].

5. Data transmission is more versatile for such signals.

### 1.1.3 Inductive Sensor Definition and Functional Block Schematic

Reducing the description area to inductive sensors, which are the subject of this book, the measurand is a mechanical state of an object (target) relative to the sensor detecting part (active face). Correspondingly, their input variables are position, level, distance, speed, acceleration, and angle, but also properties of the target, like material, thickness, surface finish, and so forth [3].

Finally, the structure in Figure 1.1 can be further particularized, as shown in Figure 1.2 [4].

The inductive sensing/sensor element (ISE/SE) of inductive sensors contains at least one inductor (coil), which is electromagnetically actuated by the metallic target. The ISE provides the record of the aforementioned variable and the conversion into a primary quantity: voltage and current, but also frequency, impedance, and so forth. Chapter 3 presents the classification and the main classes of inductive sensors and their sensing elements. The operating mode of every sensor type is then available.

The evaluation electronics (EE) is responsible for the following signal conditioning and processing up to the delivery of the output signal(s), and also for energy management, protection functions, adjustment and calibration, sensor intelligence and networking, and so forth.

## 1.2 Types of Inductive Sensors and Specific Classification Criteria

General but also some specific classification criteria are used for inductive sensors. The present section describes their specific benchmarks.

### 1.2.1 Contact or Contactless Detection

The term contact, as applied to a sensor, can be interpreted in different ways. There are essentially two adjectives which can describe contact: electrical and mechanical.

According to the block diagram in Figure 1.2, the electrical contact between the sensor and application (right side) is a sine qua non requirement, at least for the sensor power supply. A wireless radio link could replace the wired output

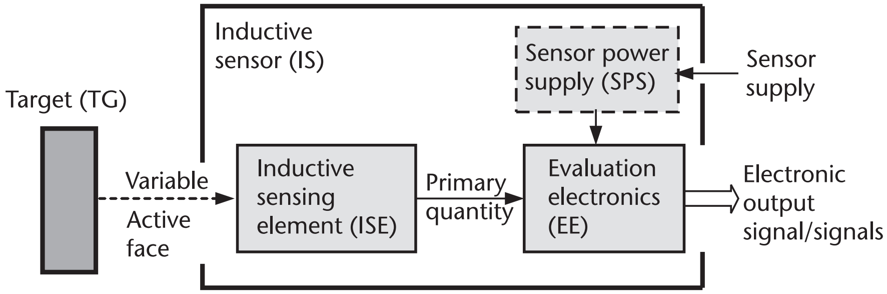

Inductive sensor (IS)
Sensor power supply (SPS)
Sensor supply
Target (TG)
Variable
Active face
Inductive sensing element (ISE)
Primary quantity
Evaluation electronics (EE)
Electronic output signal/signals

**Figure 1.2** Functional block schematic of an inductive sensor.

connection to the application. In conclusion, the electrical contact to the sensor does exist anyway.

In contrast, there are three variants generated by different possibilities with regard to the sensor input (left area of the block diagram in Figure 1.2).

If the object to be detected has a mechanical linkage with a sensor constructive part, and if this wiper arm moves an electrical contact to an internal contact-position sensitive element, that is contact sensing. A typical example is the potentiometer. The system is simple and inexpensive, but has a limited lifetime due to the friction contact. Such embodiments are not relevant for inductive sensors.

If the mechanical linkage part remains, but the coupling with the sensor element is contactless (inductive, magnetic, etc.), the method represents contactless sensing. For instance, the linear variable differential transformer (LVDT) systems described in Chapter 3 are typical contactless inductive sensors. They have a magnetic core that is rigidly coupled with the target. This moves inside a coil, which belongs the sensor element, and electromagnetically influences its inductance. Ignoring the mechanical stress, the system has high lifetime.

The third and most advantageous level is contactless actuation. Here, there is no mechanical or electrical contact between the object to be detected and the sensor. Thus, the object acts essentially as a target. Two cases are possible:

- If the object self-proximately influences the sensing element, it is called a noncooperative target. Classic examples are the very popular contactless inductive proximity sensors (IPS).
- If the object carries a small additional part, improving or allowing the detection (permanent magnet, electrical resonant circuit, etc.), the sensor works with a cooperative target.

As one might expect, contactless sensors are preferred in industrial applications. Inductive sensing technology is, by its very nature, ideal in this context.

### 1.2.2 Position versus Displacement

Unfortunately, some sensor manufacturers or technical publications confuse the terms position and displacement sensors.

Position detection (distance detection) refers to the recording of the spacing situation of the target relating to a reference datum (one end, a face, a mark on the body of the sensor, etc.); the position coordinates or the distance to the target will be measured.

In contrast, a displacement sensor is able to provide momentary distance information between the present target position and the position recorded previously, but also provides dynamic quantities like speed, acceleration, and more. Displacement sensors generally have a linear output characteristic.

### 1.2.3 Absolute or Incremental Reading

This property characterizes all measuring systems and is of course relevant to inductive sensors as well.

The absolute reading sensor provides the real output value with respect to the aforementioned datum. The major advantage of these devices is the nonvolatile character of the output information. If normal sensor operation resumes after a power interruption or after a disturbance by a strong burst of electromagnetic interference, the correct reading will be restored.

By contrast, an incremental reading sensor indicates changes in the measurand only as they occur. At any given moment, the previous, last recorded reading will be incremented or decremented in an internal sensor counter, depending on the increasing or decreasing of the input variable. Unfortunately, the output information is volatile and cannot be restored after the perturbations described above. Nevertheless, incremental systems are very popular because they can provide readings with high resolution, even over large measuring ranges.

### 1.2.4 Linear or Angular Configuration

The basic theory of operation of the linear and angular sensors is the same and they both use the same technology. The only difference lies in the kind of target movement and the corresponding topology of the sensor active face (see Figure 1.2).

## 1.3 Main Features of Inductive Sensors: Definitions and Typical Values

The specification of a sensor contains in general a large number of parameters that characterize the conditions for use, but also metrological features provided by the system. Logically, the specification of inductive sensors is somewhat different from those that describe sensor classes based on other physical active principles.

The conditions for use are generally accepted, and refer to power supply requirements, temperature ranges (storage, operating, etc.), environment demands, and so on. These parameters are presented in detail in Chapter 2.

The inductive sensor specific parameters belong to a list of needs and features which are important in describing the capability and the performance of the system. They are the subject of this section.

### 1.3.1 Supply Conditions and Limitations

For compatibility with supply networks in industrial applications, and for operator protection, the conventional sensors generally have a direct current (DC) power supply with voltages between 10 and 55V (typically 24V DC). In addition, in the U.S. market, sensors with 120V alternating current (AC) power supplies are available. The limits of the real operation voltage range, their tolerances, and the ripple limitation are product specific and will be covered later in Chapter 2.

One very important limitation is the sensor current consumption. In order to have eco-friendly systems with minimal energy consumption and to reduce the power dissipation, which decreases the internal sensor temperature and electronics drift and stress, the current demand should be below 100 mA. High-performance systems show values below 10 mA. The present electronic technologies easily allow the fulfilment of this requirement.

### 1.3.2 Sensing Range, Zero and Span, and Hysteresis

A quantitative statement about the capability of a sensor to provide a reading of the measurand can be done with the aid of the following parameters:

- Full-scale range (FSR) showing the extent of the measurand range which can be recorded: from zero to full scale for unipolar ranges, and from negative limit to positive limit for ± bipolar ranges;
- Full-range output (FRO) giving the output signal extension, which corresponds to the FSR;
- Zero-measurand and span represent another possibility, namely the lowest reading (in case it is not equal to zero) and the difference between the full-scale and this zero reading, respectively;
- Hysteresis of the output signal, defined as the difference between its values when the same measurand (e.g., position) is approached from both directions. The approaches can be unilateral (from zero to the desired value, or from full-scale to this value) or bilateral, crossing through the significant measurand value. The hysteresis may include mechanical backlash, magnetic remanence in the sensor element, and plastic parts deformation, among other factors.

The qualitative analysis is based on the major sensor notion, namely the sensor characteristic curve, which represents the dependency of the sensor output signal on the measurand (see Figure 1.3).

The main prerequisite for unambiguous detection of the measurand is the existence of a monotonic segment on this characteristic curve. This segment delimits the sensing range (SR), or measuring, working range.

Furthermore, it is highly advantageous if there is a proportionality between the measurand and the output over the entire sensing range, or at least over a large part of it (see Figure 1.3). The model-specific spread of this linearity range (LR), as

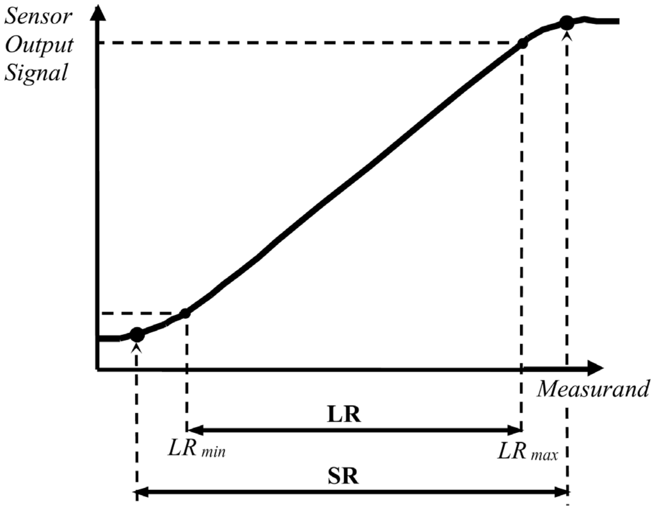


**Figure 1.3** Sensor characteristic curve showing the LR and SR.

well as its limits $LR_{\min}$ and $LR_{\max}$ and the corresponding excursion of the output, are important sensor data.

The sensor characteristic curve in the linear range corresponds to an ideal straight line having a determined slope.

### 1.3.3 Sensitivity and Nonlinearity, Linearity Error

Naturally enough, the sensor user expects from the sensor—as far as possible—an exact direct proportionality between the input variable $x$ and the output variable $y$. The geometrical illustration of this transfer characteristic, also called sensor calibration curve, should be as linear as possible.

The sensor sensitivity, or scale factor, is the slope of this curve. In general, the sensors are calibrated in such a manner that the analytical dependence $y = f(x)$ (equation of this curve) is one of three equations given in Table 1.1.

Sensors having a high and, if possible, constant sensitivity are desirable products. The constancy refers to the input value $x_a$ along the calibration curve, but also over temperature (temperature drift) and time (time stability).

The other very important sensor feature, namely the nonlinearity, describes the closeness between the output characteristic curve and a specified straight line. The deviation of the output characteristic curve from this straight line represents the linearity error $\varepsilon$, and can be expressed either in absolute values or in relative values regarding the FRO:

$$\varepsilon^{\%} = \frac{\varepsilon}{y_{\max}} \cdot 100 \tag{1.1}$$

where $y_{\max}$ is the full-range output corresponding to the high limit of the linearity range ($LR_{\max}$ in Figure 1.3).

For different reasons (including commercial), several straight lines can be considered. The usual types are diagrammatically summarized in Figure 1.4:

1. *Independent nonlinearity.* The straight line (best straight line (BSL)) is defined for positive error equal to the negative deviation, according to the least-squares criterion. This method is very popular on the market because it offers the best quality. On the other hand, the BSL nonlinearity can easily be calculated. If the straight line is represented as $y = mx + b$, the slope $m$, is found by:

**Table 1.1 Current Equations of the Sensor Calibration Curve and the Appropriate Mathematical Terms of the Sensor Sensitivity**

| *$y = f(x)$* | *Sensitivity Defined by $S(x_a) = dy/dx$ @ $x = x_a$* |
|---|---|
| $y = k \cdot x$ | $S = k$ |
| $y = k \cdot x + b$ | $S = k$ |
| $y = k \cdot x^2 + b$ | $S = 2 \cdot k \cdot x$ |

$$m = \frac{\sum_1^n x_d \cdot y_d}{\sum_1^n x_d^2} \tag{1.2}$$

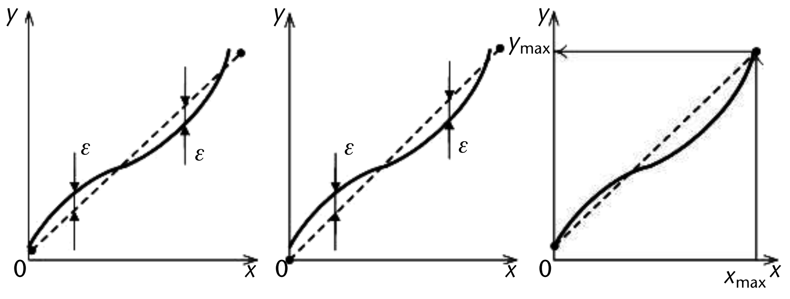
where $x_d$ and $y_d$ are the data from the measurand and output, respectively, and $n$ is the number of data points.

2. *Zero-based nonlinearity.* This procedure is similar to the independent nonlinearity, but the straight line is forced to pass through zero. This nonlinearity may be specified when wanting to ensure that the output indicates zero when the measurand is zero. The $y$-axis intercept is now called zero error.
3. *Endpoints nonlinearity.* The straight line is defined by the real output values when the input achieves the minimum and the maximum of the linearity range, respectively. Sensor manufacturers prefer not to use this method, however, because the magnitude of nonlinearity is higher. It could increase to two times the number obtained by one of the other methods.

This straight line represents the ideal case and is typically known for a certain sensor type. The sensor characteristic curve is calibrated at the end of the production process in such a way that it approaches the ideal straight line as closely as possible. Despite this, the cumulative effects of the tolerances of all sensor components result in each sensor having a sample-specific real characteristic, which deviates from the ideal straight line within a permitted range.

As a general rule, the deviation of the real characteristic curve can be quantified by adding three error values (see Figure 1.5):

1. The offset error represents a parallel shift towards the ideal straight line, through the addition of an undesired constant value to the sensor output signal.
2. The change in the ambient temperature of the sensor has a significant effect on the linearity error. The sensitivity error shows up in a rotation of the already shifted ideal straight line around a point, and results in a change of the actual sensitivity value.
3. The linearity error characterizes the slight changes in sensitivity between adjacent points, and is expressed by the deviations of the real characteristic



(a) (b) (c)

**Figure 1.4** Usual definitions for the nonlinearity: (a) Independent nonlinearity, (b) zero-based nonlinearity, (c) end-points nonlinearity.

referenced to a trend straight line. Several mathematically definable straight lines can be used for the trend straight line. In Figure 1.5, the straight line through the start and end point of the linearity range was used as a trend straight line.

The total error resulting as the sum of these three errors corresponds to the most conservative assessment of the nonlinearity. Here, the differences between real and ideal for significant points on the characteristic curve are calculated and tabulated. This procedure is known technically as “determination of the absolute linearity error.”

### 1.3.4 Accuracy, Resolution, and Repeatability: Three Precision Criterions

To qualify how exactly a sensor performs its conversion, a variety of metrological characteristics are used. The most important ones are accuracy, repeatability, and resolution.

Similar to the sensor nonlinearity definition, there are also many different, sometimes very divergent, definitions and interpretations regarding the comprehensive term precision. In our opinion, the best statement about sensor precision can be made using three mutually independent values (see Figure 1.6):

1. *Accuracy*, which corresponds to the distance between the nominal value of the output (the real value, that is, the center of the inner circle in Figure 1.6) and the actual value (statistical mean of the provided readings). It will be evaluated as the difference between the arithmetic mean value measured by a high precision reference standard system and the arithmetic mean value of the sensor output when measured repeatedly.

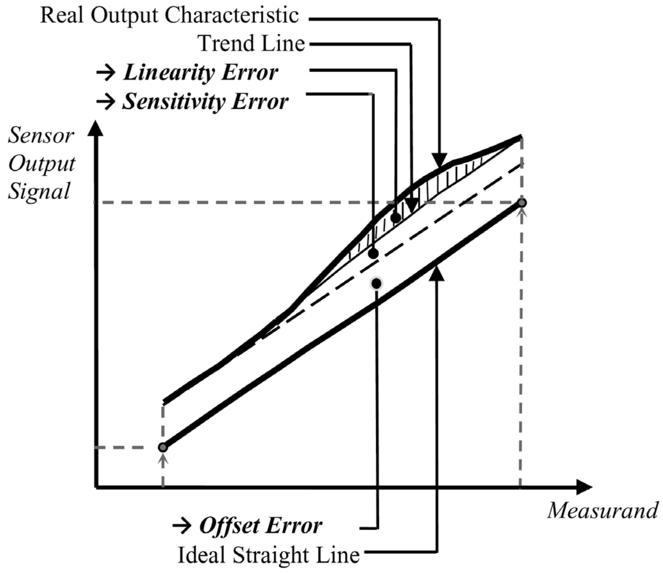


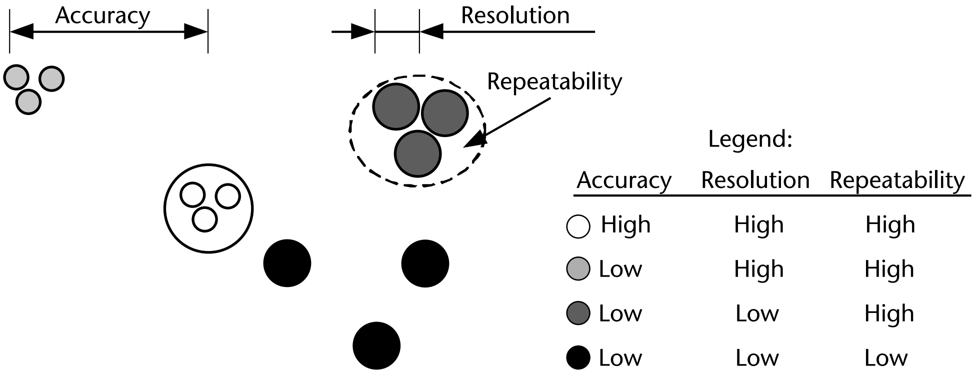
Figure 1.5 Breakdown of the total error into its three components: offset error, sensitivity error, and linearity error, respectively.

2. *Resolution*, which reflects the smallest possible change a sensor can detect in the measurand. This property is expressed by the resolution limit. A sensor with high resolution has a small resolution limit. For sensors having a digital output, the resolution limit is determined by dividing the sensing range by $2^N$, where $N$ is the bit number of the analog-digital conversion (see Section 9.8.2). Sensors with an analog output have no theoretical resolution limit. The practical value of the resolution limit is higher, and is essentially limited by the noise on the output. Inherent noise from the electronic components can be minimized, but never eliminated. The electrical noise is overlaid on the output signal and causes a high-frequency ripple of this signal. Signal changes which are smaller than the noise can no longer be clearly attributed to the wanted signal or the noise. The resolution limit can therefore never be less than the peak-to-peak value of the noise signal.
3. *Repeatability*, which is used to qualify the dispersion (sample deviation) of the provided output values. It corresponds to the scattering of the measured values, which occurs for repeated measurements for a constant value of the measurand and under fixed environmental conditions.

### 1.3.5 Drift, Temperature Effects, and Temperature Ranges

Drift encompasses the changes in the sensor output that occur even though there are no changes in the measurand or environmental conditions (constant temperature, humidity, power supply voltage, load impedance, etc.). The only variable is the elapsed time. Depending on its magnitude, the drift is described in two components: the short-time drift, which occurs in less than 24 hours and is specified as deviation in percent of FRO per hour, and long-term drift, which is specified in the same way, but the time period is a month.

The short-time drift is caused by noise, electronics, and mechanical instabilities, whereas long-term drift originates in changes of electrical components and mechanical wear or fatigue.



Legend:

| Accuracy | Resolution | Repeatability |
|---|---|---|
| ○ High | High | High |
| ◍ Low | High | High |
| ◉ Low | Low | High |
| ● Low | Low | Low |

**Figure 1.6** Graphical explanation of the three evaluation criteria for precision: accuracy, resolution, and repeatability.

When developing a new sensor, it makes sense to separate the sources for zero-drift and span-drift, and to ameliorate those accordingly by applying electrical and mechanical improvements.

These undesirable changes occur under constant ambient temperature and internal sensor temperature. In practice, it is impossible to keep these factors constant, and the aforementioned drifts will be dominated by the thermal drift of the sensor performance. The optimization of the temperature behavior and the resulting reduction in the temperature drift is a major goal of sensor design. In the author's experience, it is one of the most important goals, but also one of the most difficult.

The operating temperature ranges are specified differently depending on how the sensor will be used:

–25 to 70°C for commercial products;
–40 to 85°C for industrial or outdoor applications;
–40 to 150°C for challenging automotive implementations.

The storage temperature refers to the temperature at which the sensor can be safely stored when it is not powered, and has a larger range than the operating temperature.

### 1.3.6 Dynamic Specification, Response Time and Cutoff Frequency, and Turn On and Turn Off Times

Sensor dynamic specification describes the sensor behavior:

- In case of an abrupt stepwise change of the measurand or;
- In case of periodical, harmonic or rectangular pulse-wise variation in time of the measurand.

The system evaluation can be made using transient and frequency response analyses, respectively (by means of tests or computer-aided simulations).

The main parameters for characterizing the dynamic sensor behavior are the response time and the cutoff frequency or bandwidth.

The response time can generally be defined as the time period between the application of a change in the measurand and the resulting final indication of this change in the sensor output. A traditional method for specifying this time is to measure the elapsed time between moments when the output reaches 10% and 90%, or 5% and 95%, over its course.

The majority of conventional sensors show a low-pass filter behavior at their output, which theoretically permits definition of more intervals over time of the output signal (see Figure 1.7):

*Lag time* ($t_1 - t_0$), which elapses until a change in the output response appears, and is the consequence of several propagation delays;

*Time constant* ($t_2 - t_1$), counted until the output reaches 63% of the FRO;

*Stabilization time* ($t_3 - t_0$), i.e., the recovery time defined as summation of the lag time and time constant multiplied by five (five, because a low-pass filter specific transfer function $\{1 - e^{-\tau}\}$ generally reaches 99.3% of its final value after a time period equal to five time constants $\tau$).

The cutoff frequency, or bandwidth, is the highest frequency of a periodic sensor actuation, caused, for example, by a rotating object, whereby the output signal does not drop by more than 3 dB, that is, to approximately 70.7% of the steady-state value.

In some cases, it is also important to look at the phase lag from the measurand input to signal output, which appears at a certain frequency. A phase lag of 10° at 1 kHz, for example, means that with the measurand varying as a sine wave of frequency 1 kHz, the output signal will be delayed by 27.7 μs compared to measurand.

The desire for large bandwidth and the associated good dynamics of the sensor (on one hand), and for low noise on the sensor output with the resulting high resolution (on the other hand), represents two contradictory demands. Theoretically, the noise is distributed evenly over a very broad frequency spectrum. Filtering the high frequencies out before the sensor output results in a lower noise level on the sensor output, and therefore in better resolution; but at the same time, the bandwidth drops, thus worsening the sensor dynamics.

A time evolution of the output, similar to that shown in Figure 1.7, could appear if the supply voltage is switched on or switched off. In order to prevent initial or final instability and oscillations of the output, gating times are implemented in the large majority of conventional sensors. They are called: turn on delay (TonD) and turn off delay (ToffD), and are graphically explained in Figure 1.8. Typical time values are $(t_2 - t_0) \leq 100$ ms and $(t_4 - t_3) \leq 10$ ms.

A short glitch of the output voltage at the beginning of the TonD is possible and allowed by standards if its duration $(t_1 - t_0) \leq 2$ ms. After an accidental or provoked interruption of the supply voltage, the TonD can be retriggered only after elapse of ToffD.

### 1.3.7 Analog, Binary or Digital, and Voltage or Current Output Types

Inductive sensors perform a wide range of different outputs, from the simple switching binary output, to very complex versions for fieldbus applications.

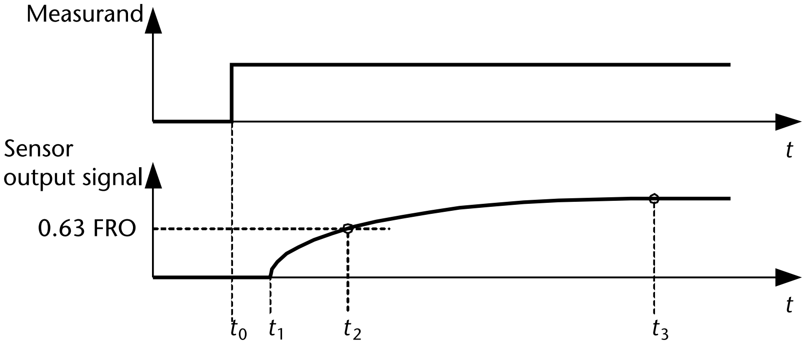


**Figure 1.7** Time response of the output signal, resulting in a step variation of the measurand.

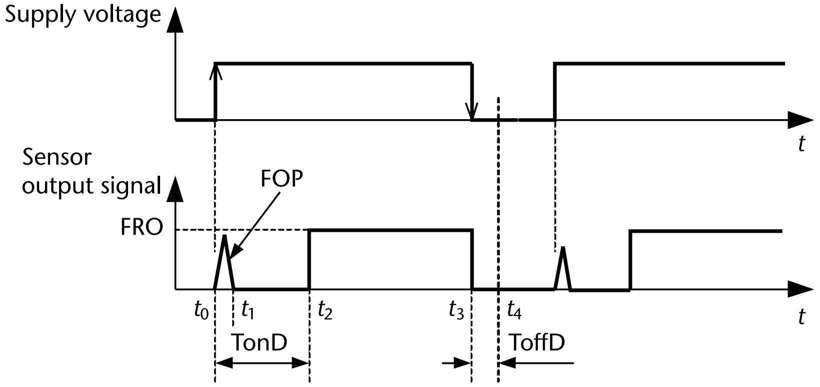


Figure 1.8 Definitions of the TonD and ToffD.

### 1.3.7.1 Binary Outputs

Binary outputs are specific to position sensors in the form of proximity sensors (PS) or other types (see Table 1.2). The connection topology consists of 2 to 4 connection lines to the application (wires).

The 3-wire version can be considered a standard (see Table 1.2). Two wires are used to connect the power supply source (battery $V_{BAT}$) to the sensor supply connections $+V_B$ and GND (ground). The third sensor connection OUT represents the switching binary output and joins the external load $R_L$ (relay, input of a programmable logic control (PLC) unit, etc.) to an internal switching driver (see in Table 1.2). Depending on the supply reference line, there are two switching output types, high-side driver (HSD) and low-side driver (LSD).

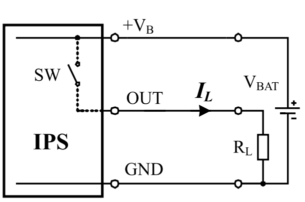
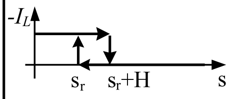
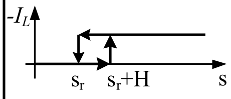
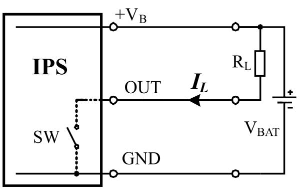
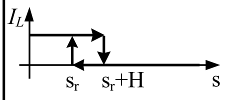
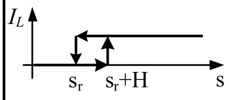
Both outputs can perform either a normally open (NO) or a normally closed (NC) switching function. Let us assume the sensor is calibrated for a switching distance $s_r$. If the actual distance between target and sensor active face is greater than $s_r$, a NO output is switched off (made inactive) and an NC output is switched on (made active). If the target approaches the sensor and passes through the switching position $s_r$, the outputs change their switching state, that is, the NO output goes ON and the NC output goes OFF.

The hysteresis described in Section 1.3.2 has a negative significance. A hysteresis $H$ with a positive property is intentionally implemented in position sensors in order to ensure the output stability. For the opposite upscale target movement, the switching process occurs at a greater distance; that is, $(s_r + H)$.

The load resistance can vary over a large range. The lowest value is limited by the maximal current capability of the switch (SW) (typically ≤ 200 mA).

The 4-wire topology is very similar. The fourth line corresponds to a second additional switching output. This can have the inverse switching function. This way the antivalent outputs increase the application safety when there is a wire break; for any distance smaller or greater than the switching distance $s_n$, there is a current flow through a line. In contrast, a second 4-wire embodiment can have the same switching function NO or NC for both outputs. The outputs are internally connected to a HSD and to a LSD, respectively. The sensor is able to drive, in parallel,

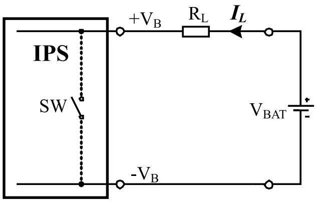
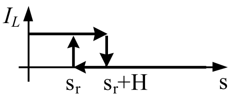
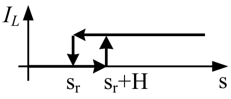
**Table 1.2** Summarizing Table for IPS: Switching Output Types in Combination with the Switching Functions [4]

| SWITCHING FUNCTION→<br>SWITCHING OUTPUT TYPE ↓ | *Make Function/ Normally Open (NO)* | *Break Function/ Normally Closed (NC)* |
|---|---|---|
| ***High Side Driver (HSD)***<br><br> +V<sub>B</sub><br>SW<br>OUT<br>$I_L$<br>V<sub>BAT</sub><br>IPS<br>R<sub>L</sub><br>GND | <br> <br>Typical application in Europe | <br> <br>Second priority in Europe |
| ***Low Side Driver (LSD)***<br><br> +V<sub>B</sub><br>IPS<br>R<sub>L</sub><br>OUT<br>$I_L$<br>SW<br>V<sub>BAT</sub><br>GND | <br> <br>Specific to American- & Asian-market | <br> <br>Lower importance |

two opposite $R_L$ loads. At the limit, the outputs can be connected together and the sensor can provide a push-pull (PP) operation.

Finally, the 2-wire version is equivalent to an electrical light switch and should belong to history by now. In fact, it is quite popular and finds universal use, especially in Japan and the Asian market. The reasons are its simplicity and lower installation costs. The application consists of a series circuit: sensor IPS, load $R_L$, and power supply source $V_{BAT}$ (see Table 1.3).

**Table 1.3** Summarizing Table for 2-Wire IPS

| SWITCHING FUNCTION → | *Make Function / Normally Open (NO)* | *Break Function / Normally Closed (NC)* |
|---|---|---|
| ***2-wire Sensor Application***<br><br> +V<sub>B</sub><br>R<sub>L</sub><br>$I_L$<br>IPS<br>SW<br>V<sub>BAT</sub><br>-V<sub>B</sub> | <br> | <br> |

If the internal SW is active (NO version is actuated, NC version is unactuated), the current flows from the battery through the load and through the sensor. The voltage drop between $+V_B$ and $-V_B$ should be as low as possible. Zero volts is impossible because the sensor electronics need a supply voltage. Thus, IPS regulates the voltage drop down to 2–3V for $I_L \leq 200$ mA (see an innovative implementation in Figure 11.5).

In the opposite case (NC version is actuated, NO version is unactuated) the internal SW is open and the current flow should be interrupted. In reality, it is necessary to maintain a residual current in order to supply the sensor with power. Competitive 2-wire sensors require such cutoff currents below 1 mA. The voltage drop on the load is very small and the sensor sees almost the entire battery voltage.

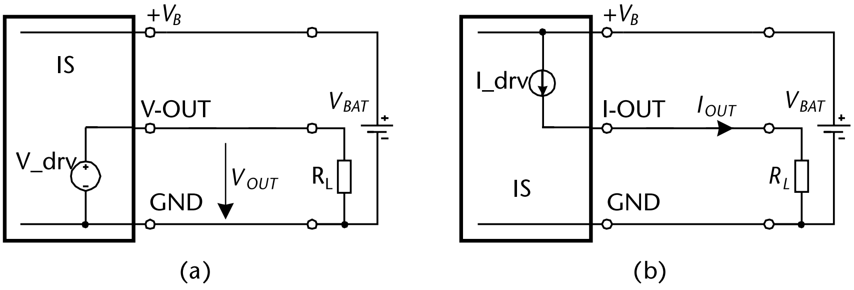
### 1.3.7.2 Analog Outputs

The analog sensor outputs are very popular and deliver a continuous signal, which is proportional to the measurand. Connection topologies with 2 to 4 wires are used most often, with the 3-wire version the most popular. There are essentially three types of analog output:

*Voltage outputs*, including 0 to 10V DC, 0 to 5V DC, and ±10V DC are fed into a high-impedance circuit or may have a load resistor, which does not sink below a specified low limit (in order to prevent an over-loading and thus false output voltage). The block schematic in Figure 1.9(a) is very similar to the HSD representation in Table 1.2. The fundamental difference is in the replacement of the internal switch SW with a measurand-controlled voltage source driver, V_drv.

*Current outputs*, including 0 to 20 mA DC, 4 to 20 mA DC, and 1 to 5 mA DC, which drive the output current $I_{OUT}$ in a loop (see Figure 1.9(b)) having a low impedance load resistor $R_L$ (zero up to a maximal resistance below 600Ω). Advantageously, by using and evaluating the loop current, the voltage drops along the connection wires are ignored. In addition, the commonly used option, 4 to 20 mA DC, allows a simple broken cable detection if the $I_{OUT}$ goes to zero.

The 2-wire current version, also called a loop-powered transmitter, has the same connection topology as a 2-wire IPS (see Table 1.3) and similar advantages.



(a) (b)

**Figure 1.9** Equivalent schematics of the (a) voltage, and (b) current outputs of an analog inductive sensor.

The switch SW is replaced by the current source I_drv. A minimum current of 4 mA is sufficient to operate the sensor. A large inductive sensor family called NAMUR has long been promoted in Europe by the NAMUR manufacturer association [5]. NAMUR ISs are used in automation networks of chemical plants.

*Ratiometric outputs* have the same typical equivalent schematic (see Figure 1.9(a)), though they provide an output voltage $V_{OUT}$ that varies as a percentage (10 to 90% or 5 to 95%) of the power supply voltage $+V_B$ to indicate the measurand value. The implementation is easier; there is no need for precise voltage reference inside of the inductive sensor. Since the sensor and the evaluation in the application refer to the same supply source, there is no additional error due to variation of the supply voltage (see Section 10.1.3).

#### 1.3.7.3 Quasi-Digital Outputs

These are often called digital outputs because they can be interfaced to digital circuits; in reality, they provide analog signals (continuous and without quantization). The popular versions are frequency modulation of an output pulse sequence, or pulse-width modulation (PWM) of the output pulses.

#### 1.3.7.4 Digital Coded Outputs

The digitalization of communication with inductive sensors is a growing development trend. For local applications of the industrial sensors, digital formats such as Serial Synchronous Interface (SSI), Serial Peripheral Interface (SPI), Universal Serial Bus (USB), and others are a solution with good cost–benefit ratio. Some serial protocols are well known; others are sensor specific and will be presented in Chapter 14.

#### 1.3.7.5 Fieldbus Compatibility

The evolution of sensor interconnectivity to complex processes is driving the massive implementation of the fieldbus solutions for both wired and wireless communication. In this context, two directions are being taken:

- Implementation of traditional fieldbus solutions: CANbus (controller area network), Profibus (application profile), and so forth;
- Development and application of sensor–actuator specific systems. The most successful: AS-Interface (actuator/sensor interface) and IO-Link (input/output link) are also described in detail in Chapter 14.

## References

[1] Dorf, R. C., *The Electrical Engineering Handbook*, Boca Raton, FL: CRC Press, 1997.

[2] Pallas-Areny, R., and J. G. Webster, *Sensors and Signal Conditioning, Second Edition*, New York, NY: Wiley, 2001.

[3] Webster, J. G., *The Measurement, Instrumentation and Sensors Handbook*, Boca Raton, FL: CRC Press, 1999.

[4] Hering, E., and G. Schönfelder (eds.), *Sensoren in Wissenschaft und Technik–Funktionsweise und Einsatzgebiete, Second Edition*, Wiesbaden, Germany: Vieweg+Teubner Verlag/Springer Fachmedien GmbH, 2018.

[5] Schiessle, E., *Industrie-Sensorik, Sensortechnik und Meßwertaufnahme, Second Edition*, Würzburg, Germany: Vogel-Verlag, 2016.

## CHAPTER 2

# Inductive Proximity Sensors: Standards, EMC/EMI, Safety, Reliability, and Availability

The goal of the first chapter was to provide preliminary basics relative to the sensors for industrial applications, in general, and regarding the inductive sensors, in particular. This information generally represents theoretical and metrological profound knowledge.

The theoretical and metrological knowledge mentioned above (see the first phrase) are necessary but not sufficient to get to know the wide field of ISs. At the same time, terms and definitions with distinctive engineering and practical character are very important. The international and/or national standards are highly suitable information sources to these ends, and they are well presented in this chapter.

The aforementioned contactless sensing and/or actuation (see Section 1.2.1) are preferred in practice and are therefore widely implemented in IS. Applications with these devices are advantageous and reliable in all application fields, beginning with the traditional industries and ending with medicine or aerospace. That is the reason for the larger volume of technical references regarding the contactless methods and products, such as standards, papers and white papers, catalogs and flyers, and others.

The key for the implementation of contactless object detections are the very popular PSs. These sensors generally have several output signals (see Section 1.3.7) and detect the target using essentially the following classical operating principles: inductive, capacitive, optoelectronic, and ultrasonic or magnetic. However, new physical working principles—especially microwave—are by now described in sensor publications or present on the market [1].

The IPS is a market leader. Due to their use in many places, some of these being very rough, the IPS has to fulfill numerous compatibility and replaceability criteria. The interoperability with several automation systems requires the fulfillment of a wide variety of standards regarding operation parameters and conditions, electromagnetic compatibility (EMC), and electromagnetic influence (EMI), CE (Communauté Européenne) mark for the European Union countries, shock and vibration, and so forth. These requirements are also described in detail in this chapter.

## 2.1 Specific Product Standards and Requirements

### 2.1.1 International Electrotechnical Commission Standard IEC 60947-5-2

The standard IEC 60947-5-2 [2] is a specific product standard and has to be primarily applied to PS and secondarily to IPS (see Figure 1.2). Additionally, the standard can be used for miscellaneous classes of inductive sensors, if these do not have a specific product standard.

This standard is the object of permanent amendments and improvements since 1992, the year of the first edition. The document cannot be considered self-sufficient. It contains an ample list of normative references. The most important are the EMC standards (series IEC 61000-4-xx), which are described in Section 2.2.

The standard has a vertical structure. It is divided in eight clauses, which contain the general definitions, classification, characteristics, product information, normal service and mounting instructions, constructional performance requirements, and test procedures to verify the features [2].

For better understanding and following of a certain characteristic, we changed the presentation logic preferring a horizontal structure, namely every characteristic is presented one time with all its aspects: definition, requirements, test conditions, and so forth.

Considering now the subclass of inductive proximity switches (IPSWs), which are basically binary convertors of the displacement information into an electrical signal, the following features are a matter of special importance for the description of this system:

- General rules and definitions;
- Operation, switching distances and actuation conditions;
- Differential travel (hysteresis);
- Switching functions and switching output types;
- Electrical characteristics;
- Mechanical specifications;
- Standard codification;
- Verification of the electromagnetic compatibility;
- Final overview of the standard specifications.

#### 2.1.1.1 General Rules and Definitions

The IEC 60947-5-2 [2] applies to:

- *Inductive* and *capacitive* proximity switches (PSWs) that sense the presence of metallic and/or nonmetallic objects;
- *Ultrasonic* PSWs that sense the presence of sound reflecting objects;
- *Photoelectric* PSWs that sense the presence of light reflecting or light absorbing objects;
- *Nonmechanical magnetic* PSWs that sense the presence of objects having magnetic properties or carrying permanent magnets.

The PSW definition complies with three essential characteristics. According to [2], the PSW:

- is self-contained (Figure 2.1);
- has a semiconductor switching element(s) whose electrical conductivity is controllable (compare with the switch symbols SW in Tables 1.2 and 1.3);
- is intended to be connected to circuits, the rated voltage of which does not exceed 250V AC or 300V DC.

Moreover, the PSW is operated without mechanical contact with the moving part: target or target carrier (see contactless actuation in Section 1.2.1).

Returning to the IPSW, this generates an electromagnetic field within a sensing zone. The geometrical intersection between the sensing zone and the sensor housing surface through which the electromagnetic field emerges is called a sensing face (see Figure 2.2). The axis perpendicular to the sensing face and passing through to its center is called reference axis (see Figure 2.3). The target movement relative to this axis can be (see Figure 2.2):

*Axial approach* if the target moves to the sensing face maintaining its reference point on the reference axis, and

*Lateral approach* if the target moves perpendicularly to the reference axis with constant distance between its surface and the sensing face.

The standard discerns between damping materials having an influence on the IPSW characteristics, and nondamping materials with a negligible influence. For the measurements of the characteristic operating distances, a specified object called standard target (see Section 2.1.1.2) is used.

The presence of a volume around the IPSW, which is kept free from any material capable of affecting the IPSW characteristics—called free zone—leads to the following classification:

*Embeddable IPSW*, when any damping material can be placed around the sensing face without affecting its characteristics. These sensor types are generally preferred because they can be mounted flush (shielded) in a metal wall (MW) or bracket (see Figure 2.2(a)).

**Figure 2.1** Different inductive proximity sensors. (Balluff GmbH/Germany.)

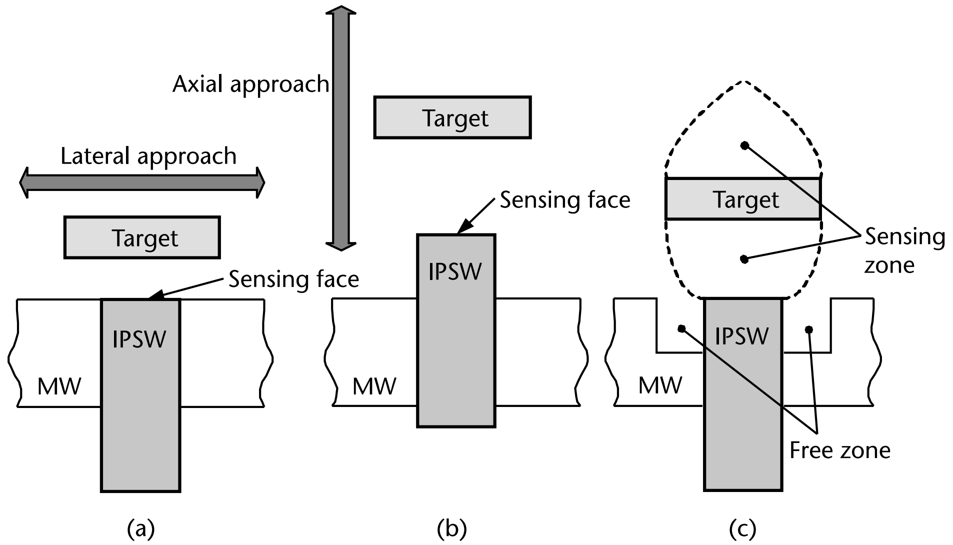


**Figure 2.2** Typical installations of an IPSW in a MW: (a) Embeddable IPSW is flush-mounted, (b) nonembeddable IPSW is nonflush-mounted (free-zone installation), (c) to reduce the MW influence, the nonembeddable IPSW is mounted in a MW cavity.

*Nonembeddable IPSW*, which requires a specified free zone in order to maintain its characteristics (see Figure 2.2(b, c)).

### 2.1.1.2 Operating Distances of the IPSW and Actuation Conditions

The IEC 60947-5-2 generally defines the operating distances ($s$) as distances where the target approaching the sensing face or moving away from the face causes the IPSW output signal to change.

Furthermore, this global quantity is more exactly specified for a standard target and its axial approach by following standard terms (see Figure 2.3):

*Rated operating distance* ($s_n$) is a conventional quantity used to designate the operating distance. This nominal value is the main feature of an IPSW, though it does not take into account manufacturing tolerances and/or variations due to external conditions such as voltage or temperature.

*Effective operating distance* ($s_r$) is the operating distance of an individual IPSW, measured at stated temperature, supply voltage, and mounting conditions (normally, not installed in a MW). This real value characterizes the sample deviation inside of a production batch and should fulfill the equation:

$$0.9\, s_n \le s_r \le 1.1\, s_n \tag{2.1}$$

*Usable operating distance* ($s_u$) is the operating distance of an individual IPS, measured under specified conditions. Independent of these conditions, its value must lie in the range of:

$$0.9\, s_r \leq s_u \leq 1.1\, s_r \tag{2.2}$$

*Assured operating distance* ($s_a$) is the distance from the sensing face within which the correct operation of the IPS is assured by taking into account all variations of the ambient conditions as well as the sample deviations in the specified ranges. Considering an IPSW without a blind zone (i.e., the zone between the sensing face and a minimum operating distance, where no object can be detected), the assured operating distance represents a worst case value and covers the range of:

$$0 \leq s_a \leq 0.9 \cdot 0.9\, s_n \tag{2.3}$$

*Repeat accuracy* ($R$) is defined as the fluctuation of the effective operating distance $s_r$ values under specified conditions. This corresponds to the metrological term repeatability (see Section 1.3.4), but is differently defined in IEC 60947-5-2 (see Table 2.1): the difference between two any measurements shall not exceed 10% of $s_r$ value:

$$R \leq 0.1\, s_r \tag{2.4}$$

The standard target in Figure 2.3 is quite specified in IEC 60947-5-2. It is a square-shaped metal plate made of Fe360-type steel, having a thickness of 1 mm (important for the eddy currents), a planned surface, and a side length $a$, which is dependent on the IPSW type. The length $a$ is equal to the greater magnitude of the following quantities:

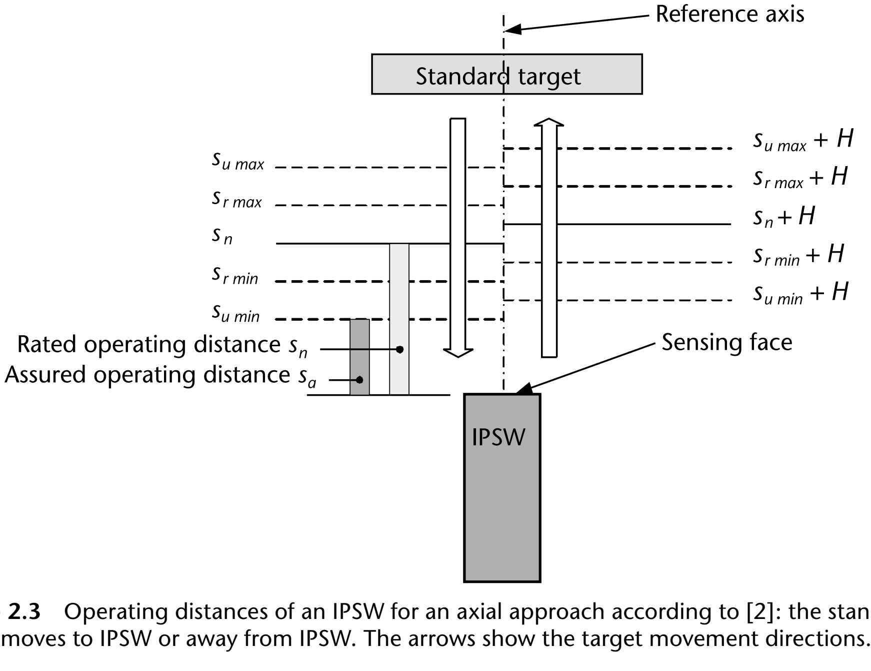

Reference axis
Standard target
$s_{u\,max}$
$s_{r\,max}$
$s_n$
$s_{r\,min}$
$s_{u\,min}$
$s_{u\,max} + H$
$s_{r\,max} + H$
$s_n + H$
$s_{r\,min} + H$
$s_{u\,min} + H$
Rated operating distance $s_n$
Assured operating distance $s_a$
Sensing face
IPSW

**Figure 2.3** Operating distances of an IPSW for an axial approach according to [2]: the standard target moves to IPSW or away from IPSW. The arrows show the target movement directions.

The diameter of the circle inscribed on the sensing face;

Three times the rated operating distance $s_n$.

### 2.1.1.3 Differential Travel (Hysteresis)

The aforementioned definitions of the specific operating distances are valid for the axial approach of the standard target (see Section 2.1.1.1). As already defined in Section 1.3.7.1, the IPSW shows greater release distances when the target is moved away in order to provide steady output states and bounce-free switching output transitions.

The differential travel ($H$) is the difference between the operating point and its corresponding release point, and is valid for all specific operating distances. According to IEC 60947-5-2, its value has to fulfill the requirement:

$$H \leq 0.2\ s_r \tag{2.5}$$

According to IEC 60947-5-2, an IPSW is subject to two test categories. The first category is the type-tests. These are once performed at the end of the design of a new sensor type (family). At least five representative samples are used. They are tested to verify compliance with the standards and to release the sensor type for serial production. The second category, namely the sampling-tests, are performed with every new manufactured sensor and are intended to verify its own quality.

The measurements of operating distances $R$ and $H$ are made in laboratory and/or production plants using calibrated measuring units, which are provided with the suitable standard target depending on the device under test (DUT). The target is moved, no faster than 1 mm per second, towards and away from the IPSW sensing face in an axial direction. General test conditions and requirements are summarized in Table 2.1.

### 2.1.1.4 Switching Functions and Switching Output Types

The definitions and classifications of the outputs in IEC 90947-5-2 are identical to the general items in Section 1.3.7.1 and do not need additional explanations.

The standard explicitly states three switching functions:

1. *Make function*, that causes the load current flow when the target enters the sensing range and vice versa if the target leaves this range;

**Table 2.1** Summary of the Working Conditions for the Testing of Operating Distances, Differential Travel ($H$), and Repeat Accuracy ($R$)

| *Parameter* | *Supply Voltage* | *Ambient Air Temperature* | *Measuring Time Period* | *Acceptance Condition* |
|---|---|---|---|---|
| $s_r$ | $U_e$ | 23°C ± 5°C | n/a | Equation (2.1) |
| $s_u$ | $0.85 \cdot U_e$ to $1.1 \cdot U_e$ | −25°C to +70°C | n/a | Equation (2.2) |
| $H$ | $U_e$ | 23°C ± 5°C | n/a | Equation (2.5) |
| $R$ | $U_e \pm 5\%$ | 23°C ± 5°C | 8 hours! | Equation (2.4) |

Note: See Section 2.1.1.5.1 for $U_e$.

2. *Break function*, which is the reversed function. Load current does not flow when the target is detected and vice versa;
3. *Make-break function* (*changeover function*), which is a combination of the switching functions stated above, namely the make function and break function, and can be implemented in sensors having two antivalent outputs (see Section 1.3.7.1).

The definitions of the switching output types in Tables 1.2 and 1.3 are in conformance with IEC 90947-5-2.

#### 2.1.1.5 Rated and Limiting Electrical Characteristics

The IEC 90947-5-2 defines and specifies a significant number of electrical characteristics, which are necessary and sufficient for industrial applications. They are listed as follows.

##### *2.1.1.5.1 Voltages*

The IEC 60947-5-2 basically defines three standard voltages:

1. *Rated operational voltage* ($U_e$), which represents the sensor supply voltage. Its value belongs to a supplier-specified range, where the high limit $U_{e\max}$ shall not exceed the values stated in the PSW definition (see Section 2.1.1.1):

   $$U_e \in \left[U_{e\min} \ldots U_{e\max}\right] \tag{2.6}$$

   The ripple voltage (peak to peak) of $U_e$ shall not exceed 0.1 $U_e$. The standard allows the sensor manufacturer to consider additionally the tolerances of the $U_e$ limits specifying an optional alternative to the $U_e$ sensor supply, namely the voltage $U_b$, having a larger range:

   $$U_b \in \left[U_{e\min} - 15\% \ldots U_{e\max} + 10\%\right] \tag{2.7}$$

2. *Voltage drop* ($U_d$) is the voltage measured at 23°C ± 5°C across the active IPSW output (see the internal SW in Figure 2.4(a)) when the rated operational output current $I_e$ flows. The units A and V in Figure 2.4(a) are ampermeter and high impedance voltmeter. The high limit of this voltage is dependent on the sensor type (DC or AC):

   $$U_d \leq 3.5 \text{ V DC} \quad \text{or} \quad U_d \leq 10 \text{ V AC r.m.s} \tag{2.8}$$

   Today's values of commercial IPSW are considerably lower, which is an important feature.
3. *Rated insulation voltage* ($U_i$) is the value of a specified voltage to which the dielectric voltage tests and creepage distances are referred.

The IPSW are hermetically sealed devices and the dielectric properties are essentially verified for every sensor family (type test) but also for each produced

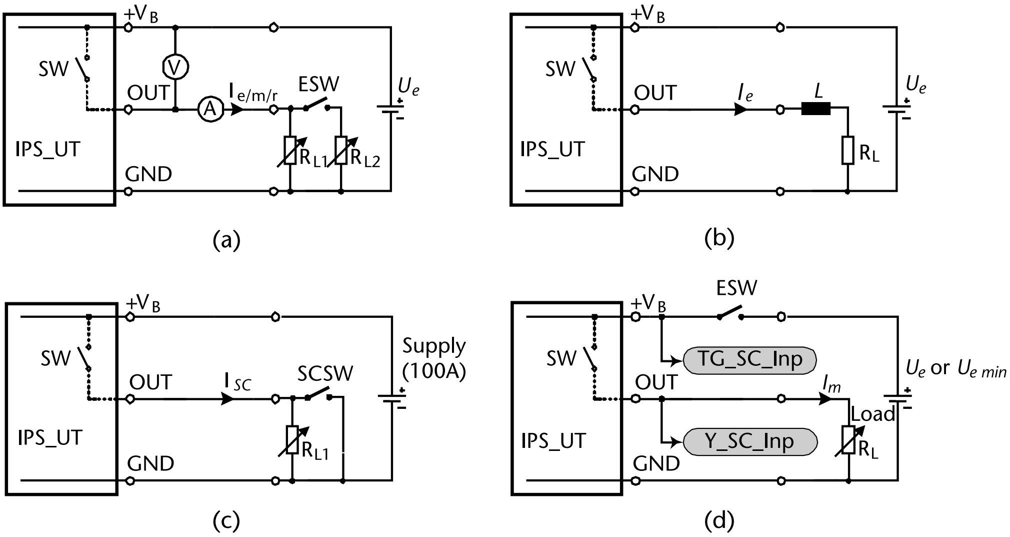

+VB
SW
OUT
IPS_UT
GND
V
A
Ie/m/r
ESW
RL1
RL2
Ue
(a)
+VB
SW
OUT
IPS_UT
GND
Ie
L
RL
Ue
(b)
+VB
SW
OUT
IPS_UT
GND
ISC
SCSW
RL1
Supply
(100A)
(c)
ESW
+VB
SW
OUT
IPS_UT
GND
TG_SC_Inp
Y_SC_Inp
Im
Load
RL
Ue or Ue min
(d)

**Figure 2.4** Basic test circuits recommended by [2] for the verification of an IPS under test: (a) IPSW_UT operation quantities, (b) making and breaking capacities, (c) short-circuit testing, and (d) time delay before availability.

exemplar (routine test). The test is to be carried out under conditions that are much closer to the actual service conditions. Surfaces of insulating sensor parts are made conducting using a covering metal foil.

The testing voltage is a sinusoidal voltage. Its root mean square (r.m.s.) value is dependent on the specified rated insulation voltage (see Table 2.2) and on the sensor type (DC or AC). This voltage is applied for 1 minute (at type test) or 1 second (at routine test) between:

Live parts of the switching element and IPSW part intended to be earthed;

Live parts of the switching element and IPSW to be touched in service (conducting or made conducting, like aforementioned).

The sensor must withstand this voltage test without unintentional disruptive discharges during the test.

**Table 2.2** The Three Levels of the Rated Insulation Voltage and the Corresponding Test Voltages According to IEC 60947-5-2

| *Rated Insulation Voltage* | | *Dielectric Test Voltage Value* |
|---|---|---|
| *DC levels* | *AC levels* | |
| 75V | 50V | 500V r.m.s. |
| 150V | 125V | 1,250V r.m.s. |
| 300V | 250V | 1,500V r.m.s. |

### *2.1.1.5.2 Currents*

The following four currents are defined in IEC 90947-5-2:

1. *Rated operational current* ($I_e$) represents the IPSW output current, which can be driven through a load (see Figure 2.4(a)). The standard values are 50 mA DC or 200 mA AC. Greater values are allowed and desired.
2. *Minimum operating current* ($I_m$) represents the current, which is necessary to maintain ON-state of the switching element SW (particular importance for 2-wire IPS). Its maximum allowed values are:

$$I_m \leq 5 \text{ mA DC or AC r.m.s, for 2-wire IPSW} \tag{2.9}$$

$$I_m \leq 1 \text{ mA DC, for 3- or 4-wire IPSW} \tag{2.10}$$

   The test jig for $I_m$ is shown in Figure 2.4(a). With the target in a position providing ON-state of the IPSW-output OUT, with supply voltage $U_e$ and with the external switch (ESW) open, the load $R_{L1}$ is adjusted to obtain the current $I_m$. The measured value shall not exceed the values in (2.9) and (2.10), respectively.
3. *OFF-state current* ($I_r$) is the current which flows through the load circuit ($R_L$ in Tables 1.2 and 1.3) of the IPSW in the OFF-state (particular importance for 2-wire IPS types). The maximum current $I_r$, which flows through the load circuit in the OFF-state, shall be:

$$I_r \leq 1.5 \text{ mA DC} \qquad \text{for 2-wire DC IPSW} \tag{2.11}$$

$$I_r \leq 3.0 \text{ mA AC r.m.s} \quad \text{for 2-wire AC IPSW} \tag{2.12}$$

$$I_r \leq 0.5 \text{ mA DC} \qquad \text{for 3- or 4-wire IPSW} \tag{2.13}$$

   The test circuit is illustrated in Figure 2.4(a). With the target in a position providing ON-state of the IPSW-output OUT, with supply voltage $U_e$ and with the ESW closed, the load $R_{L2}$ is adjusted to obtain the current $I_e$. Then the target is moved to switch the output in OFF-state. The measured current value shall not exceed the values specified in (2.11) to (2.13).
4. *No-load supply current* ($I_0$) is the current consumption from the supply source (battery $V_{BAT}$ in Table 1.2) with very high impedance of the output load. The value is stated by the IPS manufacturer. It should be as low as possible (compare with Section 1.3.1).

### *2.1.1.5.3 Switching Times, Snap (Independent) Action*

Response time of an IPS is defined as time required the switching element SW to respond after the target enters or exits the sensing zone. It corresponds essentially to the time period, which is explained in detail in Section 1.3.6.

*Time delay before availability* ($t_v$) is the start-up time between the switching on of the supply voltage and the instant at which the IPSW becomes ready to operate correctly. The time $t_v$ is equivalent to the TonD in Figure 1.8. According to the standard, $t_v$ shall not exceed 300 ms. The length of the allowed false pulse (time period $t_1 - t_0$ in Figure 1.8)—if any—shall not exceed 2 ms; its waveform is not defined.

The test jig to measure the IPSW switching behavior is represented in the simplified diagram in Figure 2.4(d). The measuring points TG-SC-Inp and Y-SC-Inp symbolize the triggering, and Y-channel scope inputs and ESW represents an external bounce-free switch. The test is performed with two target positions at $s = 1/3\ s_n$ and $s = 3\ s_n$ to assure stable ON- or OFF-states of the IPSW, respectively.

The *snap (independent) action* has to be checked at the maximum and minimum of the load current and of the operating voltage, respectively. The load in Figure 2.4(d) is an inductance-free resistor. The target is moved to provoke ON-OFF and OFF-ON transitions of the sensor output and this is observed on the scope. The switching behaviors shall be substantially independent from the target moving velocity and shall occur without oscillating or holding at any intermediate level.

#### 2.1.1.5.4 Frequency of Operating Cycles (f)

The frequency of operating cycles ($f$) is the number of operating cycles performed by the IPSW during a specified time period.

The measurement of this frequency occurs in a specific testing unit [2] with a rotating isolator disc having a diameter in accordance with the rated operating distance $s_n$ of the IPSW-UT. Its circumference is provided with metal plates identical to the aforementioned standard target (see Section 2.1.1.2) and having a distance between each other equal to $2a$. The IPSW is fixed in front of these plates at a distance of $s_n/2$. The rotation speed of the disc is increased from zero and the IPSW output signal is recorded. The output signal has a periodical time dependence, which approximates a rectangular wave with a ratio $\tau_1:\tau_2 = 1:2$, wherein the times $\tau_1$ and $\tau_2$ are the pulse width and interval width, respectively.

The rated value of the operating frequency is determined at the speed when $\tau_1$ or $\tau_2$ attain 50 $\mu$s or the output signal amplitude goes below the −3 dB level. The instant time values will be measured and used in (2.14) to calculate the switching frequency:

$$f = \frac{1}{\tau_1 + \tau_2} \tag{2.14}$$

#### 2.1.1.5.5 Normal and Abnormal Load Conditions, and Protection Functions

Normal load conditions refer to the switching capacity of the IPSW and depend on its specified utilization category. The standard defines four utilization categories—two AC and two DC—that correspond to specific applications. For example, DC-12 is suitable for resistive loads; in contrast, DC-13 is for electromagnets.

For the confirmation of the utilization categories, the standard requires to perform switching test sequences with stated parameters: supply voltage, load current, duration of ON- and OFF-state, and number of switching cycles. For example, the popular DC-12 type shell complies with the normal load test with a proper operation after 6,050 actuation cycles under the following conditions: $U_e$, $I_e$, and 1 ms transition time from OFF to ON. The test jig for the compliance is illustrated in

Figure 2.4(b). The test conditions for AC categories state a cos$\varphi$-dephasing in the load circuit. Thus, the load needs an additional inductive component.

The standard limits the external temperature rise of the IPSW due to the internal power dissipation under normal operation and load conditions at 50K, measured on the sensor enclose. The test shall be done at $U_e$ and $I_e$ and with a 2-m long connection cable, for a long time until the temperature stabilizes.

In addition, the standard refers to abnormal operation conditions (destructive or nondestructive) and to several procedures to be implemented in order to protect the IPSW (protection functions).

For the protection against short-circuit at the IPSW output, the specified rated conditional short-circuit current is 100A prospective.

The IPSW shall satisfactorily withstand (no damage) the test performed with the test jig shown in Figure 2.4(c). The IPSW having an internal, integrated short-circuit protection at its output is mounted in a similar way as in service, supplied by 1.1 $U_e$ and maintained switched ON by the corresponding target position. The normal load current is adjusted to $I_e$.

The test is performed in three successive steps. The time intervals between steps should be larger than 3 minutes (recovery time). At the beginning of a new step, the SCSW switch is closed to provoke a short circuit at the sensor output. The switch is kept closed until the internal short-circuit protection of the sensor reacts and switches off the protected output. The output of the IPSW shall withstand the stresses resulting from the short-circuit current $I_{SC}$. After the test, the normal operation and the operating distances shall remain within normal limits.

The standard does not refer to popular protections like reverse polarity protection, breaking wire protection, and so forth. By now, these are state of the art in the industry and will be described in Chapters 10 and 11.

Finally, it is important to notice that ICE 60947-5-2 also specifies service and transport conditions. The most important are the ambient temperature of –25°C to 70°C, relative humidity of the air below 90% at 20°C and the pollution degree of level 3 [2] for the application.

### 2.1.1.6 Mechanical Specifications

According to the base definition of the PSW (see Section 2.1.1.1), these systems are self-contained. In addition, annex A of IEC 60947-5-2 defines body forms and sizes. For IPSW are stated four classical embodiments (see the codification in Figure 2.5):

- *Model IA*—IPSW in a cylindrical enclosure with threaded barrel (metallic or nonmetallic), embeddable or nonembeddable, having external diameters in the range of 8 mm to 30 mm and lengths up to 100 mm. Standard thread sizes are M8×1, M12×1, M18×1 and M30×1 (Codification: M for metric, thread diameter × thread pitch in mm);
- *Model IB*—IPSW in a cylindrical enclosure with smooth metallic barrel, embeddable, having the smallest external diameter in the range 3 mm to 6.5 mm and lengths below 60 mm. Traditional diameters are 4 mm and 6.5 mm.

The trend for sensor miniaturization leads to extreme small IPSW-types (3 mm external diameter and 30 mm length) [3];

- *Model IC*—IPSW in a rectangular nonmetallic enclosure with square cross section, embeddable or nonembeddable. Standard cross sections are 26 mm × 26 mm, 30 mm × 30 mm, 40 mm × 40 mm;
- *Model ID*—IPSW in a rectangular nonmetallic enclosure with rectangular cross section and nonembeddable. Standard enclosure dimensions are 120 mm × 60 mm × 40 mm and 135 mm × 80 mm × 40 mm.

Physical dimensions, installation conditions, and corresponding rated operating distances $s_n$ of these standard types are given in annex A of IEC 60947-5-2.

For the connection with the application, the IPSWs could have connecting cable leads (typical length is 2m) with color-coded wires: brown for $+V_B$, blue for GND and black for OUT (Figure 2.4). For 4-wire version, the NO output is black and the NC output is white). However, the majority of IPSWs are standard plug-in male connectors with standard pin diagrams [2].

### 2.1.1.7 Standard Codification of the Proximity Sensors

In parallel and independent of particular product keys defined by sensor suppliers, the international standard IEC 60947-5-2 offers a universal eight-digit codification of the PSW, which is graphically summarized in Figure 2.5.

For example, according to this codification, the part number code of the most popular inductive proximity sensor, the inductive proximity switch embeddable in damping material wall mounted in a metal housing M12, with a NO-switching function and a 3-wire HSD-output type and with pin-and-socket connector is *I1A12AP2*.

### 2.1.1.8 Verification of the Electromagnetic Compatibility

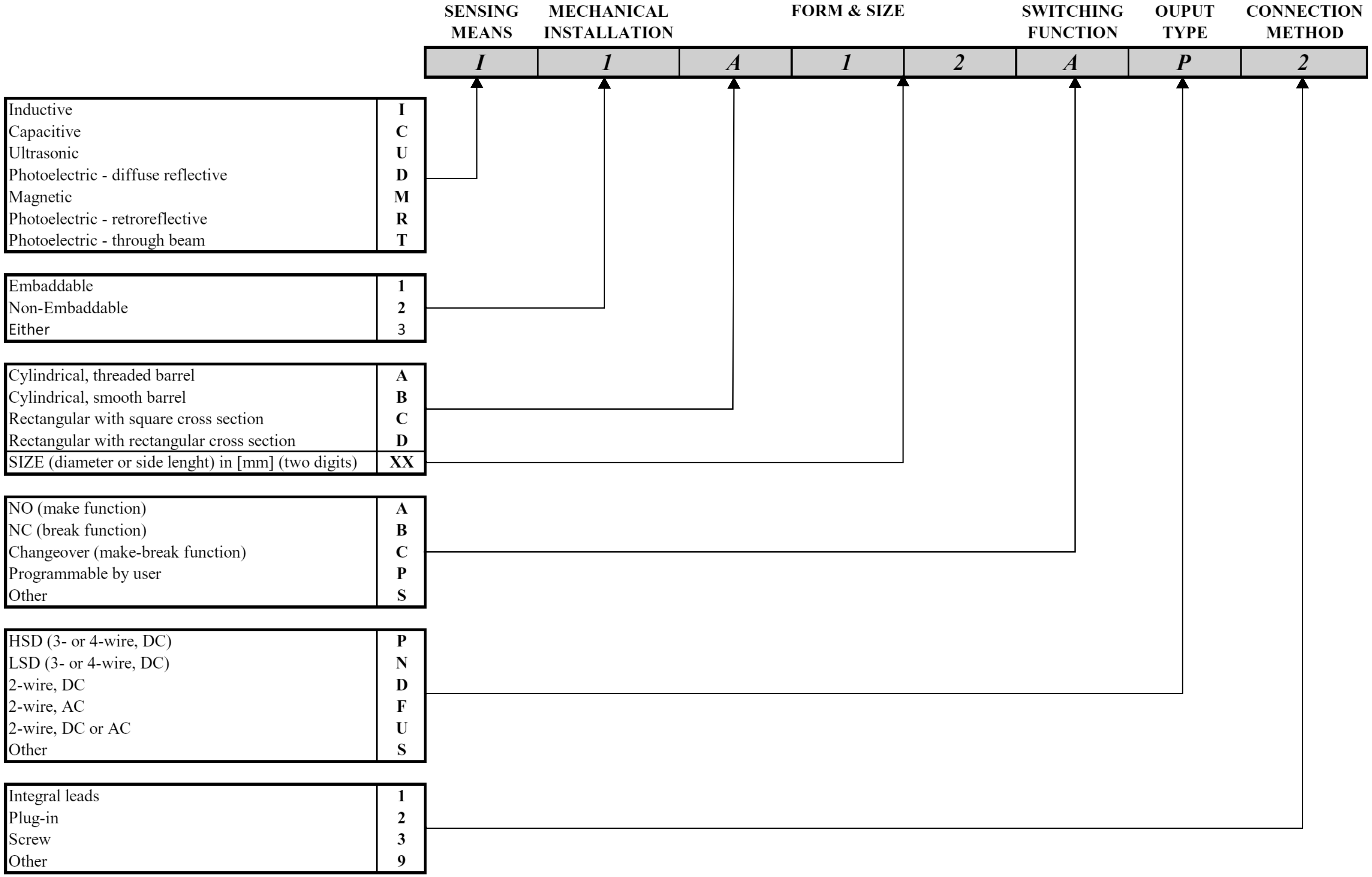
The standard IEC 60947-5-2 initially demanded to perform at least a minimal set of EMC-tests under the following conditions:

- The IPS is mounted, nonembedded, in free air and operates at $U_e$ and $I_e$;
- The EMC tests are performed with target to sensor distances of $s = 1/3\ s_n$ and $s = 3\ s_n$, which provide safe ON- and OFF-states of IPS, respectively.

The aforementioned package of EMC tests comprises:

- Electrostatic discharge immunity;
- Electromagnetic field immunity;
- Fast transient immunity;
- Impulse voltage immunity.

The tests are performed according to Sections 2.2.1–2.2.4, and to corresponding EMC standards.


SENSING MEANS
MECHANICAL INSTALLATION
FORM & SIZE
SWITCHING FUNCTION
OUPUT TYPE
CONNECTION METHOD
I
1
A
1
2
A
P
2

| | |
|---|---|
| Inductive | I |
| Capacitive | C |
| Ultrasonic | U |
| Photoelectric - diffuse reflective | D |
| Magnetic | M |
| Photoelectric - retroreflective | R |
| Photoelectric - through beam | T |

| | |
|---|---|
| Embaddable | 1 |
| Non-Embaddable | 2 |
| Either | 3 |

| | |
|---|---|
| Cylindrical, threaded barrel | A |
| Cylindrical, smooth barrel | B |
| Rectangular with square cross section | C |
| Rectangular with rectangular cross section | D |
| SIZE (diameter or side lenght) in [mm] (two digits) | XX |

| | |
|---|---|
| NO (make function) | A |
| NC (break function) | B |
| Changeover (make-break function) | C |
| Programmable by user | P |
| Other | S |

| | |
|---|---|
| HSD (3- or 4-wire, DC) | P |
| LSD (3- or 4-wire, DC) | N |
| 2-wire, DC | D |
| 2-wire, AC | F |
| 2-wire, DC or AC | U |
| Other | S |

| | |
|---|---|
| Integral leads | 1 |
| Plug-in | 2 |
| Screw | 3 |
| Other | 9 |

**Figure 2.5** Eight-digit codification of the proximity sensors according to IEC 60947-5-2.

#### 2.1.1.9 Final Overview of the Standard Specifications

Finishing the systematic presentation and completion of IEC 60947-5-2, this section offers to the reader a benchmarking in table form (see Table 2.3) of the most important IPSW characteristics. De facto, these characteristics were continuously improved in the last 20 years. Thus, the table presents both standard but also market performance features.

The number of annexes in IEC 60947-5-2 was continuously increased. Annexes B (isolation class II), C and D (additional demands for the sensor cable and connector, respectively), E (additional requirements for proximity switches suitable for use in strong magnetic fields), F (schematic PS symbols), and so forth, can be found in [2].

### 2.1.2 International Electrotechnical Commission Standard IEC 60947-5-7

The IEC 60947-5-7 [4] is a subsequent extension of the IEC 60947-5-2. Effectively, this new standard IEC 60947-5-7 modifies the relevant specifications of IEC 60947-5-2 to make these requirements suitable for proximity sensors with analogue outputs, or analog proximity sensors (APSs). In addition, the IEC 60947-5-7 can be used also for various classes of inductive sensors (see Figure 1.2) if these do not have a specific product standard.

**Table 2.3** Summary Table of the Most Important Features of the IPSW

| *IPSW Feature* | *Standard Values and Ranges* | *Market Usual, Performance Values, Ranges* |
|---|---|---|
| Rated operating distance ($s_n$) / Model IA (Section 2.1.1.6) | 2 mm–5 mm | 8 mm–20 mm |
| Differential Travel ($H$) (Hysteresis) | ≤0.2 $s_r$ | ≤0.1 $s_r$ |
| Repeat accuracy ($R$) | ≤0.1 $s_r$ @ 23 ± 5°C | ≤0.1 $s_r$ @ 23 ± 5°C |
| Rated operational voltage ($U_e$) | ≤300V DC and ≤250V AC | 10V–30V … 55V DC;<br>20V–250V AC; 20V–50V AC/DC |
| Voltage drop ($U_d$) | ≤8V/10V DC/AC @ 2-wires;<br>≤3.5V DC @ 3- and 4-wires | ≤2.5V/6V DC/AC @ 2-wires<br>≤2V DC @ 3- and 4-wires |
| Rated operational current ($I_e$) | 50 mA DC<br>200 mA AC r.m.s. | 100 or 200 mA DC<br>200 mA AC r.m.s. |
| Minimum operational current ($I_m$) | ≤5 mA DC/AC @ 2-wires | ≤3 mA DC/AC @ 2-wires |
| OFF-state current ($I_r$) | ≤0.5 mA DC @ 3- and 4-wires | ≤0.1 mA DC @ 3- and 4-wires |
| No-load supply current ($I_0$) | ≤5 mA DC @ 3- and 4-wires | ≤5 mA DC @ 3- and 4-wires |
| Time delay before availability ($t_v$) | ≤300 ms | ≤10 ms–20 ms |
| Frequency of operating cycles ($f$) | ≤1,000 Hz | ≤5,000 Hz |
| False pulse length | ≤2 ms | ≤1 ms |
| Working ambient temperature | −25°C to +70°C | −40°C to +125°C |
| Water purging protection (IP classes) (Section 2.4) | IP65 | IP67, IP68, and IP69K |

An inductive APS is defined as a device producing an output signal, which varies continuously and unequivocally depending on the distance $s$ between the target object and the sensing face of the APS. These analog devices may now consist of one or more constructive units.

IEC 60947-5-7 defines the distance/output characteristic as a relationship in the steady state of the output signal (voltage or current) with the distance $s$ (compare with Figure 1.3). Its characteristic ranges are (see Figure 2.6):

*Distance range*, extended between a lower distance (*LwD*) and an upper distance (*UpD*). Below *LwD* or above *UpD* there is no guarantee for continuous and/or unequivocal variation of the output;

*Output signal range* comprises all output values between and including two limits: lower limit (*LwL*) corresponding to the *LwD* and upper limit (*UpL*) corresponding to the *UpD*. A low limit *LwL* equal to zero is called true zero and an *LwL* unequal to zero is called live zero.

According to IEC 60947-5-7, APS are not necessarily linear devices. However, the APSs usually have a linear distance/output characteristic and their distance ranges correspond to the metrological linearity range (*LR* in Figure 1.3).

The output signal of an APS can be:

- A voltage having one of the two standard ranges:
  - *LwL* = 1V to *UpL* = +5V, or
  - *LwL* = 0 to *UpL* = +10V.
- A current having the following standard ranges:
  - *LwL* = 0 to *UpL* = 20 mA, or
  - *LwL* = 4 mA to *UpL* = 20 mA.

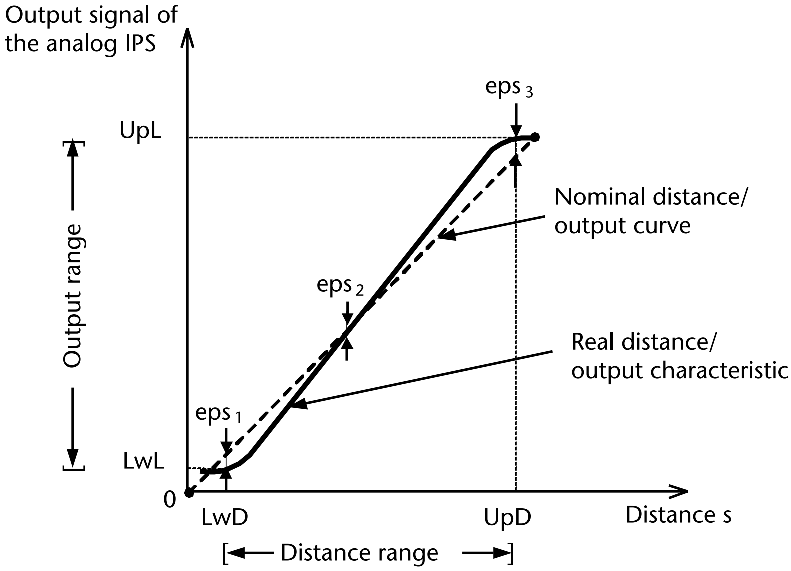

Output signal of the analog IPS
eps 3
UpL
Nominal distance/ output curve
Output range
eps 2
Real distance/ output characteristic
eps 1
LwL
0
LwD
UpD
Distance s
Distance range

**Figure 2.6** Standard definition of the distance/output characteristic of an inductive APS (IEC 60947-5-7).

The manufacturer shall state the maximum ripple content, that is, the ratio between the peak-to-peak value of the AC noise component of the output signal and the *UpL*.

The codification in Figure 2.5 still remains valid; hence, APSs are designated by a capital A placed in the fifth position.

The precision of the distance-output signal conversion (compare with Section 1.3.4) is stated in the IEC 60947-5-7 by means of the following two features:

1. *Conformity* is defined as the maximum deviation, including the manufacturing tolerances, between the nominal distance/output characteristic and measured values at defined distances. This parameter shall be within ±10% of the *UpL*.

   To measure the conformity, the APS shall be mounted as specified and the target shall be moved towards and away from the APS sensing face in an axial direction (axial approaching). The test points (at least five distances) shall be equidistant over the distance range. At least three full distance traverses in both directions shall be made.

   The conformity computation occurs in four steps. In the first step, error values are calculated as the difference between the recorded output $V_r$ and its corresponding nominal value $V_n$ for both movement directions (upscale or downscale), for every travel number $j$ ($j = 1 \dots m$), and in every test point $i$ ($i = 1 \dots n$), respectively. For example, the error for the upscale movement, travel $j$, and test point $i$ is expressed by:

$$c_{U\_j\_i} = \frac{V_r - V_n}{UpL} \cdot 100\%, \quad j = 1 \ldots m \text{ and } i = 1 \ldots n \tag{2.15}$$

   In the second step, the upscale mean errors and downscale mean errors of every travel are calculated:

$$C_{U\_j} = \frac{1}{n}\sum_{i=1}^{n} c_{U\_j\_i} \text{ and } C_{D\_j} = \frac{1}{n}\sum_{1}^{n} c_{D\_j\_i}, \; j = 1 \ldots m \tag{2.16}$$

   The third step calculates the mean values of the upscale errors and downscale errors, respectively:

$$C_U = \frac{1}{m}\sum_{j=1}^{m} C_{U\_j} \text{ and } C_D = \frac{1}{m}\sum_{j=1}^{m} C_{D_j} \tag{2.17}$$

   Finally, the arithmetic mean of upscale and downscale errors $C_U$ and $C_D$ gives the average error C:

$$C = \frac{(C_U + C_D)}{2} \tag{2.18}$$

   The conformity value is given by the highest value of $C_U$, $C_D$ and C.

2. *Repeat accuracy* represents the scattering of the output signal under specified conditions expressed as a percentage of the *UpL*.

   The repeat accuracy shall be measured at *LwD*, *UpD*, and the median value of the distance range ($eps_1$, $eps_2$, and $eps_3$ in Figure 2.6). It shall not exceed 5% of the *UpL*.

The EMC requirements for inductive PSW in IEC 60947-5-2 are also applicable for inductive APS with the following relaxation: after removing the test influence, the output signal shall return to its nominal value with the stated conformity, within a recovery time specified by the manufacturer.

The stated features in IEC 60947-5-7 represent a base specification of the inductive APS. In practice, manufacturers make a larger characterization, including additional parameters.

## 2.2 Basic and Specific EMC/EMI Standards

Section 2.1 presented the standard requirements stated in IEC 60947-5-2 that the inductive sensors have to meet to be suitable for industrial applications. One component part refers to the electromagnetic compatibility (see Section 2.1.1.8). The standard simply stated which the necessary EMC tests are, without explanations how to perform them. For the performance of these tests, the standard references the basic and specific EMC standards and specifies only the levels, which are to be achieved. For this reason, the present section dives into the EMC world.

In connection with the increasing complexity and sensitivity of sensor electronic components, the susceptibility to external influences comes into consideration (see Figure 2.7).

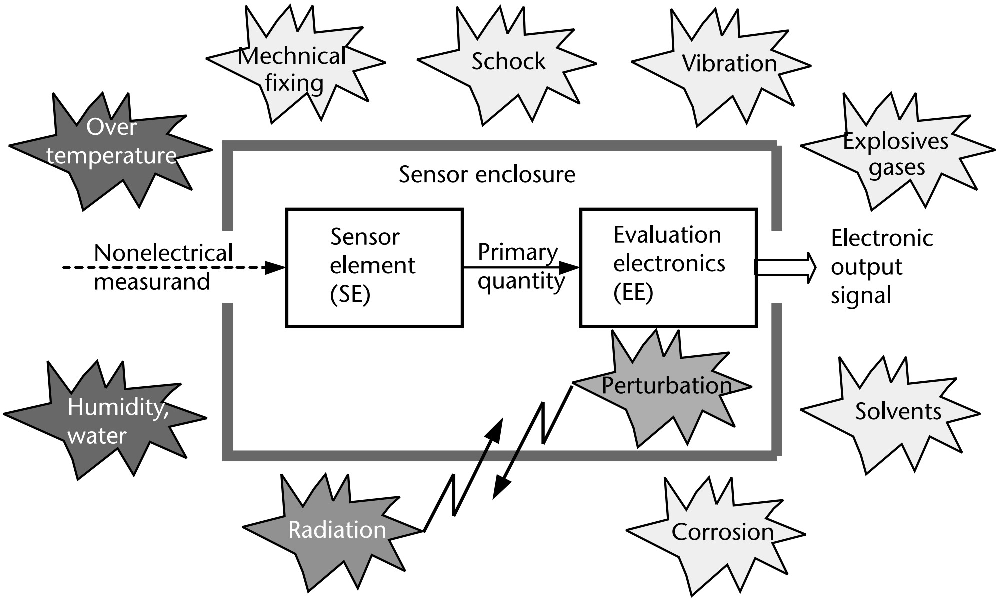

Mechnical fixing
Schock
Vibration
Over temperature
Explosives gases
Sensor enclosure
Nonelectrical measurand
Sensor element (SE)
Primary quantity
Evaluation electronics (EE)
Electronic output signal
Perturbation
Humidity, water
Solvents
Radiation
Corrosion

**Figure 2.7** Various sensor-disturbing sources.

An important class is the electromagnetic disturbances. The sensor has to provide an undisturbed output under these intensified perturbations. Either they can be volitional perturbations (power networks, broadcast emissions, radar, etc.) or unintentional, which can be generated by various sources, from atmospheric up to cosmic radiations.

The EMC of a device is a rating of how well it will operate in an environment of electromagnetic radiation. When electromagnetic radiation emitted from someone's device affects the performance of a sensor, it is called EMI.

EMC includes two key aspects: device susceptibility (immission) and device emission. A sensor, which fulfills the EMC requirements, will not be affected (susceptibility) by EMI emitted from other devices at reasonable levels, nor will it generate sufficient perturbing electromagnetic energy (emission) to disturb the performance of other devices.

Due to their operating principle, the susceptibility has a major importance for the inductive sensors. In contrast, they emit low-level energies so that their emission has a minor significance.

The way of looking at the EMC problem is well regulated by a large number of EMC standards. This section describes the common EMC requirements, in general, and the specific EMC requirements for inductive sensors, in particular.

The EMC/EMI disclosure can be classified in three categories:

1. *International basic EMC standards*. They describe the measuring procedure and the requirements of every particular EMC test. In addition, they contain threshold values in the form of severity levels, which are addressed by the next inferior standard categories. The identification of the basic EMC standards is IEC 61000-4-xx, where xx are numeric characters starting with 00.
2. *Product and product family standards*. Important products have separate standards that state functional but also EMC requirements. That is the case of the product standard IEC 60947-5-2 for proximity sensors (Section 2.1.1). The standard concretely specifies the needed EMC requirements for PSW (Section 2.1.1.8) and the corresponding severity levels (highlighted in Table 2.4).
3. *Generic EMC standards*. These standards are based on the general EMC standards. They specify the requirements for products and systems operating in residential or industrial environments (see Table 2.5) and apply to

**Table 2.4** Testing Voltages and Field Strengths (±10%) for EMC Examinations

| | *ESD Test Voltage* | | *RFI* | *EFT Test Voltage* | | *Surge* | *CMCD* | *MFI* |
|---|---|---|---|---|---|---|---|---|
| *Severity Level (SL)* | *Contact Discharge* | *Air Discharge* | *Field Strength* | *Coupling Network* | *Coupling Clamp (CCC)* | *emf Value* | *r.m.s. Value* | *Continuous Field* |
| 1 | 2 kV | 2 kV | 1 V/m | 0.5 kV | 0.25 kV | 0.5 kV | 1V | 1 A/m |
| 2 | 4 kV | 4 kV | 3 V/m | 1 kV | 0.5 kV | 1 kV | 3V | 3 A/m |
| 3 | 6 kV | 8 kV | 10 V/m | 2 kV | 1 kV | 2 kV | 10V | 10 A/m |
| 4 | 8 kV | 15 kV | n/a | 4 kV | 2 kV | 4 kV | n/a | 30 A/m |

**Table 2.5** Generic EMC Emission and Immunity Standards [5]

| *International Standard IEC* | *European Standard EN* | *Topic* | *Residential, Commercial and Light Industrial Environments* | *Industrial Environments* |
|---|---|---|---|---|
| 61000-6-3 | 50081-1 | Emission standard for | X | |
| 61000-6-4 | 50081-2 | Emission standard for | | X |
| 61000-6-1 | 50082-1 | Immunity standard for | X | |
| 61000-6-2 | 50082-2 | Immunity standard for | | X |

products for which there is no specific product standard or there is a product standard however without specified EMC requirements.

The sensor test conditions (actuation mode, power supply or mechanical installation, etc.) are very firmly stated in the standards. Sensor damages and/or destructions due to the EMC tests are not allowed.

Finally, the test result, namely the behavior of the sensor during and after the test stimulus, is defined by one of following acceptance criteria (AcC):

- *Acceptance criterion A:* no failure and no false output pulses (FOP) (see the short glitch over the time $t_1 - t_0$ in Figure 1.8) during the test;
- *Acceptance criterion B:* allowed FOP (shorter than 2 ms) or no safe operation during the stimulus, but full operation after the test within a specified recovery time;
- *Acceptance criterion C:* allowed FOP longer than 2 ms during the test or deviant behaviour that needs a system reset or—at the limit—nonreversible falling out.

For successful usage in industrial applications, the following six EMI immunity properties are normally required:

1. Resistance to electrostatic discharge (ESD);
2. Resistance to radiated electromagnetic fields (radio frequency interference (RFI));
3. Resistance to electrical fast transients (EFT);
4. Resistance to lightning surges;
5. Resistance to common mode conducted disturbances (CMCD);
6. Resistance to power frequency magnetic field immunity (MFI).

The operated characteristics of the inductive sensors shall be maintained at all levels of EMI up to the severity level required by the applicable standards or greater level stated by the manufacturer.

For the emission, the standards to meet are generally considered the requirements for the CE Mark. Since 1996, all products that are liable to cause, or be affected by EMI must have the CE mark if they will be sold in the European Union countries. For the purposes of compliance testing, electromagnetic energy may either be conducted or be radiated through space.

In order to make sure that the sensor meets the EMC requirements, the tests take place in two steps at two locations:

1. *In-house test laboratory*. Initial tests are made to verify with design samples if the final product will pass the real test;
2. *Third party* (not the developer, not the customer), namely test laboratories that are certified to perform EMC test and to confirm by written reports the fulfillment of the industrial EMC requirements.

### 2.2.1 Resilience against Electrostatic Discharges

The ESD is a high-voltage discharge similar to the one which appears when a person walks across a carpet in a chamber with low air humidity, stores electrostatic charges and touches an isolated metal part like a sensor, for example. Sensors must be able to survive for discharge voltages up to specified levels.

The definition of the resistance to electrostatic discharges, the test conditions, and the requirement to certify this important property are described in the basic EMC standard IEC 61000-4-2.

ESD is simulated by applying the specified voltage directly to metal parts of the IPS under test (IPS_UT) or through an air gap that simulates an approach of the human operator (see Figure 2.8). A number of 10 single pulses having a typical duration 60 ns and 1s periodicity shall be applied for every test mode (contact or air gap). The voltage levels of the ESD test impulses, corresponding to the four standard severity levels and two different discharge arts, are summarized in Table 2.4. It results a discharge current with an amplitude of more amps.

### 2.2.2 Resilience against Radiated Electromagnetic Fields

The resistance to RFIs describes the property of an electronic device to withstand electromagnetic perturbations (in a frequency range below 3 GHz), generated by radio emitters, mobile phone, radar stations, or other closed range appliances.

The definition of the resistance to external radiated electromagnetic fields, the test conditions, and the requirements to certify this immunity are described in the basic EMC standard IEC 61000-4-3.

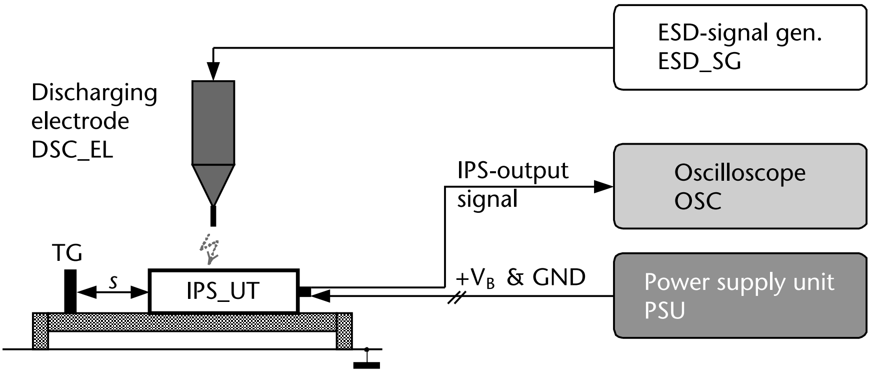


**Figure 2.8** Schematic representation of the test jig for the ESD immunity test

RFI is tested placing the IPS_UT in a gigahertz transverse electromagnetic cell (GTEM) chamber (see Figure 2.9) and exposing it to an anechoic RF field (the chamber walls are provided with electromagnetic absorbers). Characteristic for this investigation are long exposure time and tests in three orthogonal directions. The field strength values are shown in Table 2.4. To generate this field, the multiband antenna in the chamber is supplied with a high power RF signal having the following parameters:

- Carrier frequency up to 2.7 GHz;
- Modulation signal: 1kHz, sinusoidal, 80% modulation depth.

### 2.2.3 Resilience against Fast Transients: EFT, Burst

The EFT test simulates the burst of high voltage and high range of frequency that can occur when relay contacts open (particularly when the closed contacts conducted high currents though inductive loads). Other burst sources are transformers, magnetic valves, and high voltage circuits.

The definition of the resistance to fast transients superposed the power supply, the test conditions (with all sensor-connecting leads placed in a capacitive coupling clamp (CCC) or wired through a coupling network), and the requirements to certify this important property are described in the EMC basic standard IEC 61000-4-4.

The test signal, which is delivered by a burst signal generator BST_SG (Figure 2.10), is periodical and consists of a cluster of burst pulses with a 15-ms cluster duration followed by a 285-ms pause. The burst pulses inside of the cluster are a dual exponential one: every pulse has an exponential rise time to the peak value of 5 ns, followed by an exponential decay back to zero over a longer time period (just above 100 ns). The burst pulses have a middle pulse duration of 50 ns and a 5 kHz recurrence frequency. It results a broadband frequency spectrum up to 500 MHz. The peak values are given in Table 2.4. The pulse coupling occurs by means of a CCC for supply and output sensor wires (see Figure 2.10).

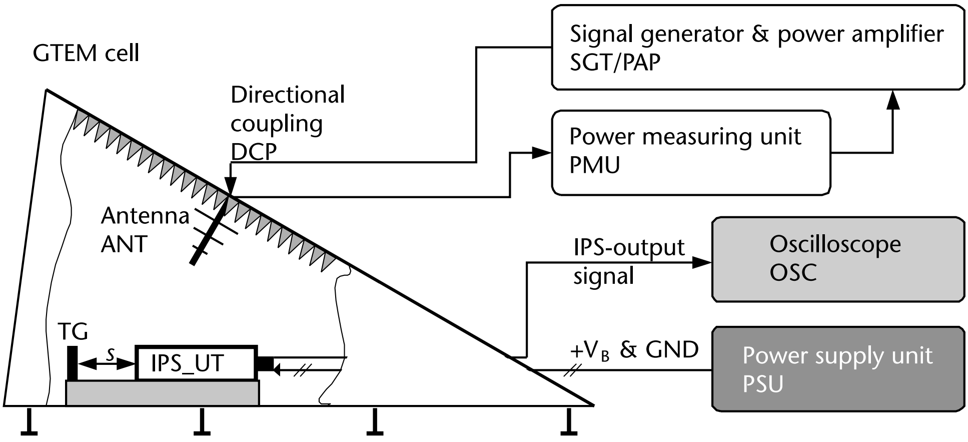


**Figure 2.9** Schematic illustration of the equipment for the RFI immunity test using an anechoic GTEM chamber.

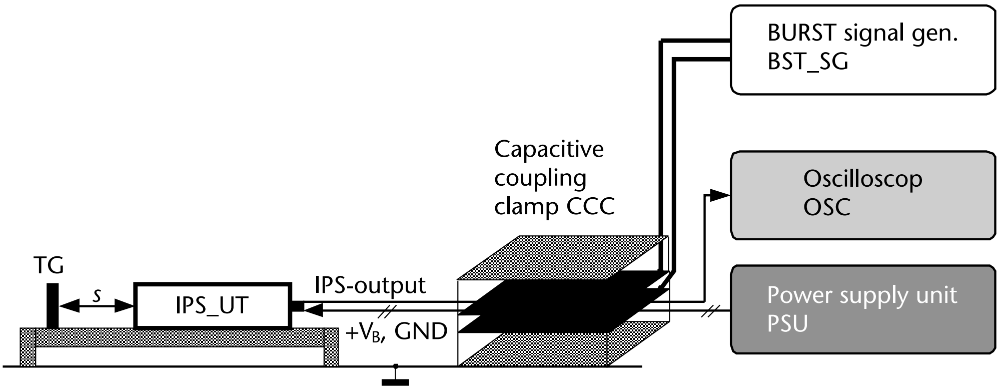


**Figure 2.10** Schematic representation of the test jig for the EFT/burst immunity test.

### 2.2.4 Resilience against Impulse Voltage (Surge)

A surge test simulates the increased voltage levels that may be induced between device conductors (supply and/or outputs) when a lightning strike occurs in the general areas or switching on/off processes in energy wide networks are performed. The strikes could induce voltages ranging from 500 to 5,000V and having a duration stretching from 1 ms to 10 μs, respectively.

The definition of the resistance to high energetic voltage pulses, the test conditions and the requirements to certify this immunity are described in the EMC basic standard IEC 61000-4-5.

The test signal is a single high-energy pulse, which is a dual exponential one: the pulse has an exponential rise time to the peak value of 1.2 μs, followed by an exponential decay back to zero over a longer period of time (see Figure 11.9(a)). It has a duration of 50 $\mu$s in the pulse middle (compare with 50 ns of the burst pulse). The peak values are given in Table 2.4.

After every test step, the sensor is normally supplied and the normal operation test (at $U_e$ and $I_e$) is performed.

According to IEC 60947-5-2, the surge test of the PSW is a sole type test and is performed with production release samples under the following specific conditions:

- The sample is not powered during this test;
- The test impulses shall be applied:
  - between all terminals connected together and the earth;
  - between terminals intended to be connected to power supply;
  - between each output terminal and each terminal intended to be connected to power supply.
- Three positive and three negative impulses shall be applied between each terminal pairs at intervals greater than 5s;
- An external power resistor of 500Ω is interconnected (see Figure 2.11) to simulate a surge generator with higher impedance (the standard surge-generator output impedances lies between 4Ω and 40Ω). Accordingly, the surge current achieves peak values of 1 to 8A (see peak voltages in Table 2.4). All internal clamping-circuits between the tested terminals of the PSW must withstand this high current without any destruction. The subject of surge protection

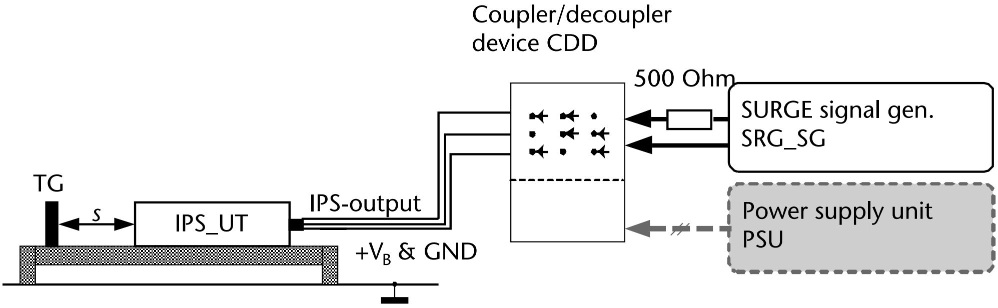

Coupler/decoupler
device CDD
500 Ohm
SURGE signal gen.
SRG_SG
Power supply unit
PSU
TG
s
IPS_UT
IPS-output
+V_B & GND

**Figure 2.11** Schematic diagram of the test jig for the surge immunity test.

(clamping devices, integrated protection electronic circuits, etc.) is continued in Chapter 11.

### 2.2.5 Resilience against Common Mode Conducted Disturbances

In contrast to the evaluation of the resistance to RFI where the immunity to radiated high-frequency (HF) fields is tested, the test CMCD proves the resilience against conducted induced HF interferences. The test checks the device behavior during continuous disturbances caused by electromagnetic fields, which are interspersed through the sensor cable.

The definition of this resistance, the test conditions and the requirements to certify this property are described in the EMC basic standard IEC 61000-4-6.

The coupling of the HF disturbance is inductive and is made by a directional coupling network (DCN) (see Figure 2.12). The signal generator provides a modulated test signal, whose amplitude values are shown in Table 2.4 and whose HF parameters are:

- Frequency range: 150 kHz to 80 MHz;
- Modulation signal: 1 kHz, sinusoidal, 80% degree of modulation.

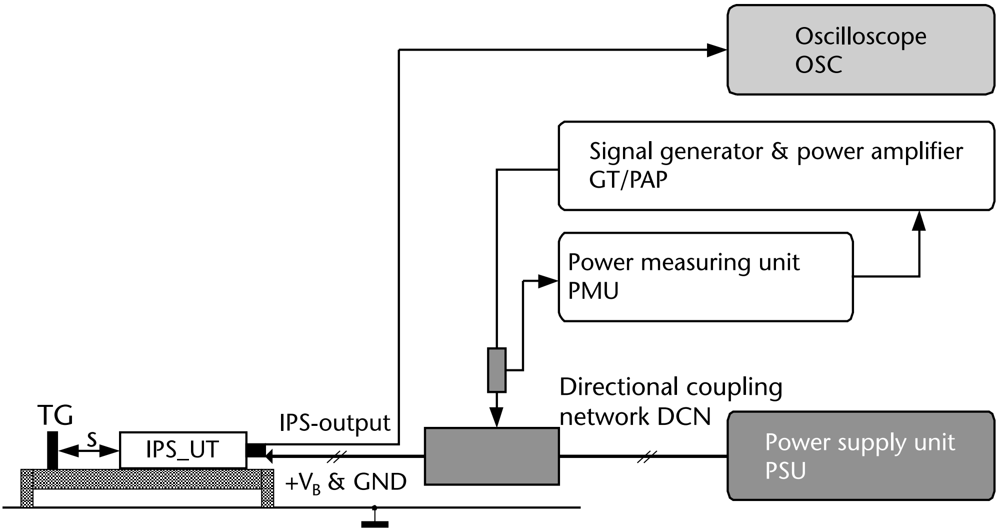

Oscilloscope
OSC
Signal generator & power amplifier
GT/PAP
Power measuring unit
PMU
Directional coupling
network DCN
Power supply unit
PSU
TG
s
IPS_UT
IPS-output
+V_B & GND

**Figure 2.12** Schematic representation of the test jig for the CMCD immunity test.

### 2.2.6 Magnetic Field Immunity Test

This test shell evaluates the immunity to external magnetic fields having the power frequency (50/60 Hz), which inherently appear in industrial environments.

The definition of this immunity, the test conditions, and the requirements to certify this important immunity are described in the EMC basic standard IEC 61000-4-8.

The test jig is an induction coil in the form of one aluminum solid winding (square shape with a 1-m side length) and supplied by a power current source. The current frequency corresponds to the regional mains power frequency (50 Hz or 60 Hz). The coil generates a magnetic field having magnitudes up to 100 A/m (constant) or 1 kA/m (temporary).

### 2.2.7 Immunity to Voltage Dips and Interruptions

The immunity to voltage dips and interruptions (VDI) test shell evaluates how the sensor can suppress fluctuations or interruptions of its supply provoked by jitters into power networks or in the load. These perturbations have a sporadic character and their magnitude, repeat frequency, and duration are very different.

The definition of this immunity, the test conditions, and the requirements to certify this property are described in the EMC basic standard IEC 61000-4-11. This states the following perturbation parameters:

Supply voltage drops: 30%, 60%, and 95%;

Power interrupt periods of the 50 Hz power supply: 10 ms, 20 ms, 100 ms, 200 ms, 500 ms and 1s.

### 2.2.8 Summary of the EMC Test Conditions for Inductive Proximity Sensors

Essentially, the basic EMC standards are valid for various electric and electronic devices; at the same time, the required levels (see Table 2.4) are device-specific.

The product standards IEC 60947-5-2 and IEC 60947-5-7 consequently specify the severity levels (SLs) for PSW and APS (highlighted in Table 2.4).

The IEC 60947-5-2 started in 1992 with four EMC requirements (see Sections 2.2.1 to 2.2.4). Over the years, the requirements list was extended (see Sections 2.2.5 to 2.2.6) and some severity levels increased (see Table 2.6).

## 2.3 Shock and Vibration Requirements

Most IPS will undergo some level of shock and vibration exposure during normal use. This section qualifies the IPS for two purposes:

1. *Performance investigation*, to see how normal shock and vibration levels affect the sensor precision (see Section 1.3.4);
2. *Survivability evaluation*, to estimate how long the operation period under greater levels of shock and vibration is before failure.

**Table 2.6** Up-to-Date Summary of the EMC Test for Inductive Proximity Sensors, Binary and Analog, According to IEC 60947-5-2 [6]

| *Resilience Against* | *Applicable Basic Standard* | *Test Conditions* | *Severity Level (SL) and Acceptance Criteria (AcC)* |
|---|---|---|---|
| ESD (Section 2.2.1) | IEC 61000-4-2 | • short-time, single electrostatic ESD-pulses: ns–range, ≤30 kV;<br>• 10 pulses/1s interval;<br>• direct (contact) and/or indirect (air-gap). | SL = 2 (4.0 kV) @ direct discharge or<br>SL = 3 (8.0 kV) @ indirect<br>AcC: B |
| RFI (Section 2.2.2) | IEC 61000-4-3 | • narrow-frequency band, scanned up to 2.7 GHz;<br>• AM-modulation, sinusoidal, 1 kHz, 80% depth;<br>• field strength ≤ 10 V/m. | SL = 2 (3 V/m)<br>AcC: A |
| EFT (Section 2.2.3) | IEC 61000-4-4 | • dual exponential, triangle shape pulses, 50 ns typ. duration @ 50% level;<br>• pulse repetition rate inside burst: 5 kHz;<br>• burst cluster: 15 ms duration, cycle period: 300 ms;<br>• capacitive coupling into sensor cable. | SL = 3 (1.0 kV for CCC)<br>AcC: B |
| Surge (Section 2.2.4) | IEC 61000-4-5 | • dual exponential, triangle shape pulses, 50 $\mu$s duration @ 50% level;<br>• single positive and negative pulses but also alternating pulses;<br>• test stimuli between respective two sensor connections;<br>• 500Ω external resistor as generator output impedance. | SL = 2 (1.0 kV)<br>AcC: A |
| CMCD (Section 2.2.5) | IEC 61000-4-6 | • frequency range: 150 kHz to 80 MHz<br>• AM-modulation, sinusoidal, 1 kHz, 80% depth;<br>• Inductive, network coupling into sensor cable. | SL = 2 (3 V r.m.s.)<br>AcC: A |
| MFI (Section 2.2.6) | IEC 61000-4-8 | • field generated by a solid induction coil in the form of one square winding (1 m side length);<br>• coil supplied by a power current source;<br>• frequency: 50/60 Hz (depending on the region). | SL = 4 (30 A/m, continuous)<br>AcC: A |

*From:* [6].

The test equipment comprises a vibration machine setup having a test fixture for the sensor, supplied by a chain: a signal generator and power amplifier, and carrying an accelerometer as a measuring part for the entire control system (see Figure 2.13).

The vibration tests are performed at several discrete frequencies and amplitudes. It is important to try to discover any resonance tendency. If any resonance frequencies are found, the sensor shall be tested at this frequency for extended time to make sure that the proper performance is maintained or—much better—the sensor design can be modified to damp this tendency.

The shock tests are performed swiping the shock-pulse parameters like shape, acceleration, duration, and additionally the number of pulses.

Similar to the EMC evaluation, initial vibration and shock tests shall be done during design at the sensor provider. After this in-house verification, additional tests are usually conducted at a certified testing lab to make sure that industry standards are met.

For the PSW and APS, the product standard IEC 60947-5-2 states the stability against vibrations and mechanical shocks as it refers to specific standards.

For the vibration resilience IPS evaluation, the tests are made according to the specific standard IEC 60068-2-6 in compliance with the following conditions:

- Vibrations applied in each positive/negative direction along three mutually orthogonal axes (six separate tests);
- Frequency range: 10 to 55 Hz;
- Amplitude: 1 mm;
- Sweeping cycle duration: 5 min;
  - Duration of endurance at resonant frequency: 30 min in each of the three axes.

For shock resilience IPS evaluation, the tests are made according to the specific standard IEC 60068-2-27 in compliance with the following conditions:

- Pulse shape: half sine;
- Peak acceleration: 30 $g_n$ (300 m/s$^2$);

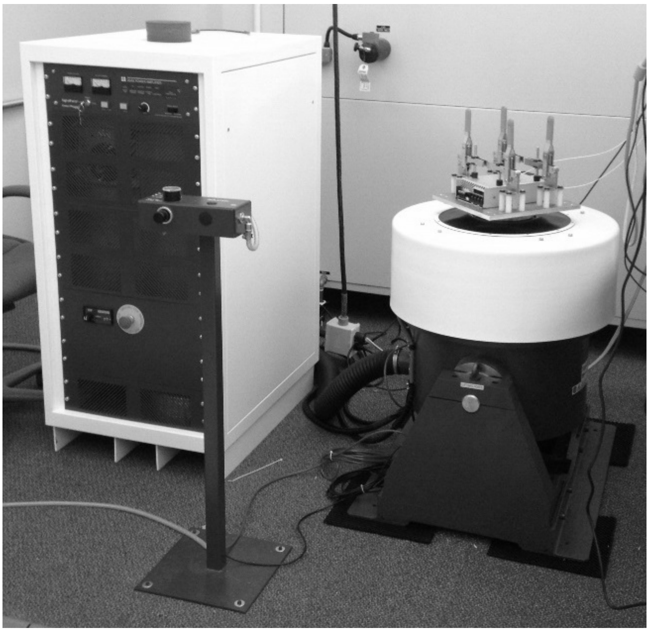


**Figure 2.13** Typical forcer for vibration tests. (Data Physics San Jose-CA.)

- Pulse duration: 11 ms;
- Six shocks applied in each direction positive/negative along three mutually orthogonal axes (six separate tests).

Following the tests, the operating characteristics of the IPS shall remain unaffected.

## 2.4 International Protection Classification

The majority of electronic devices must stand rough conditions and work for their specified lifetime. Nevertheless, they are exposed to dirt, dust and humidity.

International protection (IP) codes (IP-codes) summarize the protection against parts contact (first IP digit) and infiltration of water and dirt (second IP digit). They define how and where the devices can be used without getting a safety risk [7].

Usually, the first digit of the IP-code with relevance for IS is five, or more commonly, six (see Characteristic 1 in Table 2.7). According to the product standard IEC 60947-5-2 for PSW, they shall have at least the degree of protection IP65. In reality, the products on the sensor market today provide IP67 to IP69K (highlighted in Table 2.7).

**Table 2.7** Summary of the IP 2-Digit Codes

| *First IP Digit* | *Characteristic 1 for Protection Against Contact* | *Second IP Digit* | *Characteristic 2 for Waterproofing* |
|---|---|---|---|
| 0 | No special protection | 0 | No waterproofing |
| 1 | Protection from solid objects greater than 50 mm in diameter | 1 | Protection from dripping water |
| 2 | Protection from object not greater than 12 mm in diameter | 2 | Protection from vertically dripping water (tilted up to 15°) |
| 3 | Protection from object not greater than 2.5 mm in diameter | 3 | Protection from sprayed water (tilted up to 60°) |
| 4 | Protection from object not greater than 1 mm in diameter | 4 | Protection from splashed water |
| 5 | Complete protection against contact,<br>Protection against dust deposit | 5 | Protection from water projected from a nozzle |
| 6 | Complete protection against contact,<br>Protection from infiltration of dust | 6 | Protection against heavy seas, or powerful jets of water. |
|  |  | 7 | Protection against immersion (1h @ 1m) |
|  |  | 8 | Protection against complete, continuous submersion in water. |
|  |  | 9K | Protection against the ingress of water from all directions even under greatly high pressure on the housing. (High-pressure/steam cleaner, 80 to 100 bar) |

## 2.5 Intrinsic Safety, Product Safety Certification

The IS must, at times, be able to operate in areas that may experience a hazardous (flammable or explosive) atmosphere intermittently or continuously.

The classical methods to prevent combustion or explosion are:

- *Intrinsic safety*, which consists of minimizing the ignition energy in the device;
- *Explosion proofing*, a method which could generate an explosion should occur. An explosion-proof housing is designed to be capable to withstand the resulting explosion;
- *Purging*, a method that removes hazardous gas from the area surrounding the sensor, replacing this gas with a nonflammable gas.

The National Electrical Code (NEC) in the United States specifies three types of hazardous locations:

- Class I (gas or vapor);
- Class II (dust);
- Class III (fibers or flyings).

In addition to the class, hazardous locations are separated into two divisions:

- *Division 1* comprises locations where hazardous conditions may exist under normal operating conditions;
- *Division 2* covers locations where hazardous conditions may exist only under abnormal operating conditions.

The *European zones 1* and *2* are similar to the divisions but not identical. In addition, there is the *European zone 0*, where the hazardous conditions exist continuously.

The elements required to initiate combustion are generally shown in the classical combustion triangle (see Figure 2.14).

Three components are needed before combustion/explosion can occur. The fuel and oxidizer (oxygen) components occasionally belong to certain processes.

Thus, the only way to prevent explosion is to make sure that sufficient energy to cause ignition is not present. That is the principle behind intrinsic safety (ISF).

The so-called *devices for EX applications* are intentionally designed to provide a high ISF level. Traditional measures implemented in ISF-rated sensors are:

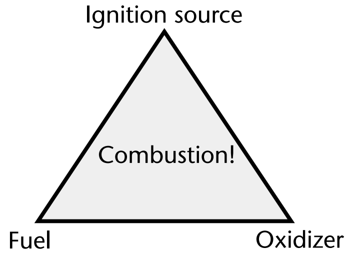


**Figure 2.14** Combustion triangle.

- No hot spots inside;
- Low enough outside case temperature;
- No stored energy in excess;
- Use barriers between hazardous area (intrinsically safe sensor) and nonhazardous area. These electronic barriers could be passive components (combination Zener diodes with resistors) or active devices (ASIC, galvanic isolators, etc.);
- Designed with ISF-approved components, except for some non-ISF-compliant devices, which do not have energy storage capability (light-emitting diodes (LEDs), switches, etc.).

Sensors intended to be used in hazardous areas must obtain an approval from a certified approval agency. A primary approval agency known in North America but also in Europe (by the European Committee for Electrotechnical Standardization) is Underwriter's Laboratories. Its ISF standard *UL 913* is accepted worldwide.

To conclude, the ISF method is the most suitable way to provide inductive sensors for EX applications. The ISF method could be implemented in parallel to the use of explosion-proof housings or purged areas. This additionally increases the system safety.

## 2.6 Reliability and Availability

### 2.6.1 Mean Time Between Failures, Mean Time to Failure, and Failure Rate and Availability

The probability that a sensor will perform and be free of failure is called reliability.

Intrinsic reliability is a function of the quality of design and is the highest theoretically possible value. Due to unexpected factors (abnormal using conditions, quality problems in the production, lack of maintenance) the reliability drops approaching the operational reliability.

To quantify the reliability, the following quantities are used:

- *Mean time between failures* (*MTBF*), which estimates service life of a sensor until it requires repair or replacement;
- *Mean time to failure* (*MTTF*), which should be used for nonrepairable devices.

The MTBF shall be applied over a long period of time to a statistically large sample.

If the loss rate of a device remains constant throughout its life, the MTBF is the inverse of the failure rate (FR). This is normally assumed, because a newly installed product may be subject to infant mortality, whereas an old product may contain parts that are wearing out. The FR is expressed using any measurement of time, but the hour is the most common unit in practice.

The inner electronics of the IPS contain a large number of semiconductor devices. The semiconductor industry particularly uses a specific FR definition, namely the failures in time (*FIT*). The FIT rate of a device is the number of failures that can be expected in one billion ($10^9$) device-hours of operation:

$$FR = \frac{FIT}{10^9} \tag{2.19}$$

There are two possibilities to determine the value of MTBF:

1. To calculate MTBF: The method is suitable for newly designed sensors. The device is broken down into all its component parts (electronic components, solder joints, connectors, etc.). Each component has an assigned individual failure rate. These values are added up and the inverse of this total failure rate gives the MTBF:

$$MTBF = \frac{1}{FR} = \frac{1}{\sum_{i=1}^{n} FR_i} \tag{2.20}$$

   where $FR_i$ is the failure rate of the $i$ component;
   $FR$ is the total failure rate of the device expressed in hours;
   $n$ is the number of components.

2. To demonstrate MTBF: Conversely, if a large number of the sensors have been installed and operating in the field for a long time, the MTBF can be determined with (2.20) but the $FR_i$ values are field failure rates.

The probability $R(T)$ that a given device having a known MTBF will perform without failure over a given period $T$ can be calculated with the following exponential formula:

$$R(T) = e^{-T/MTBF} \tag{2.21}$$

Usually, the sensor suppliers guarantee sensor operation for $T = 10$ years (about 90,000 hours) and the reality confirms this affirmation.

Correspondingly, assuming a probability of $R(T) = 0.85$ (85%), the design shall provide a MTBF value:

$$MTBF = -\frac{T}{\ln R(T)} \tag{2.22}$$

The result, MTBF ≈ 553,800 hours, reflects the high quality of the design and fabrication of the sensors for long-term industrial applications.

Finally, the sensor user would like to know another important aspect of the reliability called availability (AV). AV is the extent to which the sensor is fit for use at the time when it is actually needed. It is expressed as:

$$AV = \frac{MTBF}{MTBF + MTTR} \tag{2.23}$$

where MTBF is the mean time before failure and MTTR is the mean time to repair.

If the sensor has a moderate MTBF and needs a long time to repair, it results in a low availability. The IS has high MTBF values and requires a short time to repair, and if this were the case, the availability can theoretically increase up to 100%.

### 2.6.2 Highly Accelerated Life Test

This test, in development since 1988, gains increased importance despite the fact that it is not just yet a standard requirement. It is a time-efficient stress testing methodology for enhancing product reliability [8]. Highly accelerated life test (HALT) is a test technique called test-to-fail, where a product is tested until failure. It is a qualitative test method and does not help to determine or demonstrate the reliability value or failure probability in field.

The main goal is to discover weak spots in the system design or fabrication by means of a small number of prototypes. This occurs by an acceleration of the aging and damaging processes caused by temperature stress, vibrations, and thermal shocks.

The test is performed in a specific HALT chamber. It consists of the following actions:

- *Stepwise test at low temperature* (first single-stress examination): The sensor's ambient temperature will be decreased starting at 20°C in steps of 10K (dwell time 10 minutes) until the specified lower operating limit and still the lower destruct limit will be achieved. These significant temperatures will be recorded.
- *Stepwise test at high temperature* (second single-stress examination): The sensor's ambient temperature will be increased starting at 20°C in steps of 10K (dwell time 10 minutes) until the specified upper operating limit and still the upper destruct limit will be achieved. These significant temperatures will be recorded.
- *Temperature cycling* (third single-stress examination): At least five cycles between lower operating limit and upper operating limit with a dwell time of 10 minutes will be done.
- *Vibration test* (fourth single-stress examination): The vibrations start at 5 G r.m.s. (root mean square acceleration) and the value will be systematically increased (increment of 5 G r.m.s. and dwell time of 10 minutes). The upper limit remains under 60 G r.m.s. The destruction limit—if any—will be determined.
- *Combined stress test* (combined stress examination): It is a final combined test consisting of a temperature cycling with more stated temperature profiles in parallel to vibration tests (to up to five vibration levels), and is limited by the destruction limit determined during the vibration tests.

Every step is a 360° process. If a destruction appears, the sample will be investigated by a fault analysis, repaired and/or improved, and retested under tightened conditions, and so on.

At the end, a result concerning the limits of the product will be generated and can be provided to the customer, if desired.

## References

[1] Fericean, F., A. Dorneich, R. Droxler, and D. Kräter, "Development of a Microwave Proximity Sensor for Industrial Applications," *Sensors Journal*, IEEE, Vol. 9, No. 7, 2009.

[2] IEC-International Electrotechnical Commission, *International Standard IEC 60947-5-2, Low-Voltage Switchgear and Controlgear, Part 5: Control Circuit Devices and Switching Elements, Section 2: Proximity Switches*, www.vde-verlag.de/iec-normen/, Germany: 2014.

[3] Jagiella, M., S. Fericean, and A. Dorneich, "Progress and Recent Realizations of Miniaturized Inductive Proximity Sensors for Automation," *Sensors Journal*, IEEE, Vol. 6, No. 6, 2006.

[4] IEC-International Electrotechnical Commission, *International Standard IEC 60947-5-7, Low-Voltage Switchgear and Controlgear, Part 5: Control Circuit Devices and Switching Elements, Section 7: Requirements for Proximity Devices with Analogue Output*, www.vde-verlag.de/iec-normen/, Germany: 2003.

[5] http://amvelectronica.com, *Notas/Estandares_emc.pdf*, Edition 5/2000.

[6] Ekbert Hering, Gert Schönfelder (Hrsg.), *Sensoren in Wissenschaft und Technik—Funktionsweise und Einsatzgebiete, 2nd. Edition*, Wiesbaden / Germany, Vieweg+Teubner Verlag / Springer Fachmedien GmbH, 2018.

[7] IEC-International Electrotechnical Commission, *International Standard IEC 60529, Degrees of protection provided by enclosures (IP Code)*, www.vde-verlag.de/iec-normen/, Germany: 2013.

[8] IEC-International Electrotechnical Commission, *International Standard IEC 62506, Methods for product accelerating testing*, www.vde-verlag.de/iec-normen/, Germany: 2014.

## CHAPTER 3

# Inductive Sensors: Definitions, Main Types, and Market Share

## 3.1 Overview of the Sensor Classification

Technical publications generally classify sensors following the sensor application field criterion. They are important for users who look for problem solvers for their application. As an example, the voluminous reference [1] amply describes such numerous sensor applications.

For technical people, there are additional, highly significant classification criteria. The first criterion evidently refers to the operating principle of the sensor (resistive, inductive, capacitive, magnetic, etc.). An attentive classification and description according to this major sorting mark is available in [2].

A second criterion has an electrotechnical substrate and assumes that the SE, which generates the primary electrical quantity (compare with Figure 1.2), is equal to an impedance $\underline{Z}$ with its component's resistance $R$ and reactance $X$:

$$\underline{Z} = R + jX \tag{3.1}$$

This general expression directly defines two large sensor categories. If the measurand influences the real part $R$ or the imaginary part $X$ of the SE, then the sensor belongs to the resistive or the reactive category.

A further, third point of view is based on an energetic consideration. Actually, for most operating principles (see the first criterion), the measurand conversion is provided either by direct energy conversion or by control of the energy flux in the sensor front-end.

The SE of sensors, which belongs to the direct energy conversion class, is able to extract nonelectrical energy from the measurand, whose value is measurand-dependent, and to convert it into electricity. The sensor is self-generating and utilizes an electrical generator principle. This energy conversion could be reversible or nonreversible.

A first sub-class exploits the behavior of electrical contacts made from two different materials: metal and semiconductor and/or isolator. They are used to achieve a temperature measurement (primarily based on the Seebeck effect) or to obtain special applications like electrophoresis, electro-osmosis, flow detection, and so forth.

Another major sub-class refers to the evaluation of mechanical quantities (distance, speed, acceleration). Popular systems are:

- *Electrostatic* convertors, which consist of a capacitor having its geometry influenced by a target. Classical technical realizations are microphones for sound, vibrations, or membrane-bending applications;
- *Piezoelectric* convertors having a piezoelectric core to record forces and speed-ups, but also to measure acoustic signals. The reverse piezoelectric effect, that is, electrostriction, is the fundament of ultrasonic emitters, including the SE of the ultrasonic PSW (see Section 2.1.1.1).
- *Electrodynamic* convertors are systems with a high relevance for the subject of this book, namely IS. They consist of a moveable inductor (coil), which is driven by the target. The coil moves in a permanent magnetic field and generates electricity that is dependent on the force or speed of the target. The system is analyzed in Section 3.3.2.
- *Electromagnetic* convertors, which essentially have an inversed topology. The coil is by now fixed in a magnetic core and the target modifies the core geometry. The system performs an immediate conversion of distances in output currents and has a high relevance for our book subject. Sections 3.3.3 and 3.3.4 describe in detail the corresponding embodiments.

In addition, the direct energy conversion class contains several sub-classes without significance for this book, listed here to only be enumerated without any details: optoelectronic, pyroelectric, electrochemical convertors, and so forth.

On the inside of sensors, which belong to the second class with control of the energy flux, the measurand is a control quantity, which modulates an auxiliary energy flux in the sensor front-end. The sensor is passive and the controlled energy could be a harmonic quantity whose parameters (amplitude, frequency and/or phase) are modulated. It can also be a DC value, which is pulse-modulated or directly influenced by the measurand. This second class covers important sensor families:

- Resistive sensors;
- Galvanomagnetic (Hall) and thermomagnetic sensors;
- Inductive sensors;
- Capacitive sensors;
- Optoelectronic sensors;
- Gas sensors;
- Electrolytic conductivity meters.

Finally, a fourth classification criterion refers to the measurand type. This classification hits every operating principle. Restricting now the multitude of sensor measurands to the five main nonelectric types (see Section 1.1.1), the representation in Figure 3.1 shows classical ways to convert these variables in an electrical quantity, which can be evaluated by the sensor EEs (see Figure 1.2). The symbols of variables and electrical quantities are explained in Table 3.1.

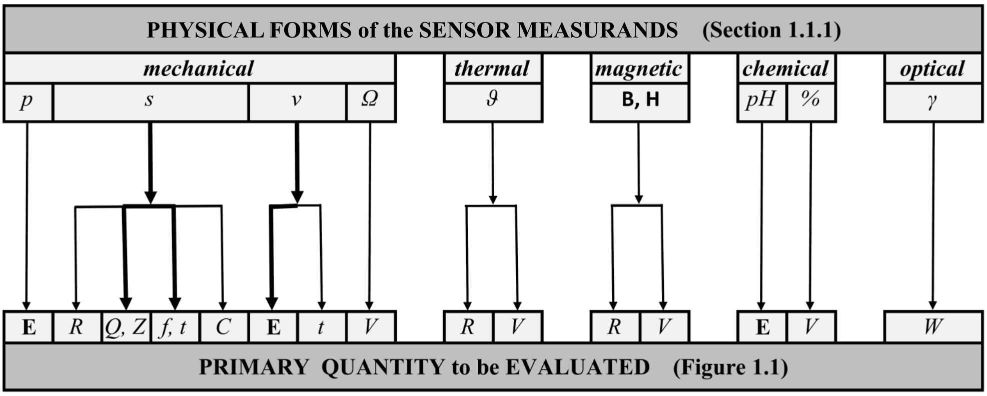

PHYSICAL FORMS of the SENSOR MEASURANDS (Section 1.1.1)
mechanical
thermal
magnetic
chemical
optical
p
s
v
Ω
ϑ
B, H
pH
%
γ
E
R
Q, Z
f, t
C
E
t
V
R
V
R
V
E
V
W
PRIMARY QUANTITY to be EVALUATED (Figure 1.1)

**Figure 3.1** Conversion paths of nonelectrical measurands into primary electrical quantities. The major conversion paths of mechanical distances (of every sort) realized in inductive sensors are highlighted.

## 3.2 Specific Embedding of Inductive Sensors

A first introductory presentation and definition of ISs has already been done in Section 1.1.3 and Figure 1.2. The present section completes this definition and localizes the IS in the classifications in Section 3.1.

ISs are primarily based on the principles of magnetic circuits. The physical definition of the IS includes this device in the reactive sensors category. In an IS occurs a direct or indirect evaluation of the reactive sensing element SE (inductive), by determining the flux changes provided by moveable targets in defined electromagnetic surroundings or by sensing SE parameter deviations because of target enter in SE self-field.

Inherently to its functional principle, the IS contains at least one inductor (coil) which has the property to induce an electromagnetic field.

**Table 3.1** Glossary for Figure 3.1

| *Physical Nonelectrical Measurand* | | *Electrical Primary Quantities* | |
|---|---|---|---|
| p | Pressure | **E** | Electric field strength |
| s | Travel, distance, position | R | Resistance |
| v | Speed | Q | Q-factor, quality factor |
| Ω | Angular frequency, rotational speed | Z | Impedance |
| ϑ | Temperature | f | Frequency |
| **B** | Magnetic flux density | t | Time |
| **H** | Magnetic field strength | C | Capacitance |
| pH | pH-value, ionic concentration | V | Voltage |
| % | Volume gas concentration | W | Electric energy |
| γ | Light quantum, photons | | |

The measure of the capacity for magnetic induction—called inductance $L$—can be generally expressed in terms of the magnetic flux $\Phi$ by the equation:

$$L = n\frac{\Phi}{i} = n^2\frac{\Phi}{V_m} = \frac{n^2}{R_m} \tag{3.2}$$

in which $n$ is the number of the coil turns, $i$ is the current flowing through the coil, $V_m$ is the magnetomotive force (mmf) (measured in units of ampere-turns) and $R_m$ is the magnetic reluctance.

The magnetic flux $\Phi$ in the core is proportional to the core permeability $\mu$ and the (3.2) can be still developed:

$$L = n\frac{\Phi}{i} = n\mu A\frac{H}{i} = n\mu_0\mu_r A\frac{H}{i} \tag{3.3}$$

where $A$ is the cross-sectional area of the core, $H$ is the absolute value of the magnetic field intensity (considered to be constant), $\mu_0$ is the permeability of free space, and $\mu_r$ is the relative permeability of the coil core (if any).

In conclusion, the inductance $L$ can be controlled if the physical measurand can influence the material, geometrical and/or electromagnetic quantities. This fact leads to division of inductive sensors in ISs based on material interaction or on geometry interaction.

Inductive sensors are mass-produced devices and are widely used in industry. They are robust and compact, and less affected by environmental factors: humidity, dust, and so forth (see Figure 2.7).

ISs are essentially utilized to convert mechanical information (every sort) regarding a certain target. The most popular conversion paths of mechanical distances (predominantly linear and angular displacements) realized by ISs are bold printed in Figure 3.1. In addition, inductive gaging is used to measure thicknesses (ranges from micrometer up to more millimeter), tilts and inclinations, shocks and vibrations, strain and forces, torques and bandings, liquid intrusions and levels, fluid flows and conductivity (noninvasive gaging), viscosity, and so forth [1].

## 3.3 Main Types of Inductive Sensors

Based on several publications that deal with ISs, small different opinions can finally be observed. Nevertheless, a leitmotif can be recognized and this is similar to the experience of the author. Essentially, seven main types of inductive systems/sensors can be identified. Their basics and characteristics are presented further.

### 3.3.1 Magnetoelastic Systems

The system belongs to the category based on material influence (see (3.3) in Section 3.2) and it is passive. It operates using a phenomenon known as the Villari effect (reverse magnetostriction) and utilizes the fact that a change in mechanical stress of a ferromagnetic material causes its permeability $\mu$ to alter. For detailed information about the magnetoelastic effect and the dependence permeability versus mechanical stress, a recommended reference is [3].

The mechanical stress can be tensile or compressive strain. Consider, for example, an SE consisting of a coil wound onto a rod core of magnetostrictive material (see Figure 3.2). The force $F$ to be measured is applied on this core, stressing it and causing its elastic deformation but also a change in its permeability. Finally, the coil impedance is a function of $F$:

$$|\underline{Z}| = \mathcal{F}(F) \tag{3.4}$$

This change can be monitored and is a rate of the applied force and of the target displacement, indirectly. To exercise a force on the core, it is necessary to have a rigid linkage between the target and this core. Consequently, magnetoelastic ISs provide a contactless sensing and not actuation (see Section 1.2.1).

Typical suitability materials are nickel or permalloy (nickel-iron alloys), but also magnetically soft ferrites having low saturation flux density (see Chapter 6).

The system is suitable only for small travels as the elongation of the elastic deformation is very limited. It has poor linearity and is subject to hysteresis (see Section 1.3.2). The advantage of the system is the wide range of bandwidth up to 15 kHz (see Section 1.3.6) and therefore the adequateness for high-speed events.

The sensitivity is dependent on the core cross-sectional area and increases when it is subjected to pure torsion, regardless of direction. As an example, [4] refers to Vitrovac amorphous magnetic compositions provided by Vacuumschmelze GmbH [5], which are able to evaluate strains up to $5{\cdot}10^8$ newtons per meter or more. The permeability value is in inversely proportional to the strain magnitude.

### 3.3.2 Electrodynamic Systems

These systems belong to the second category based on geometry influence (see (3.3) in Section 3.2) and have a large application area. It consists of the inductive perception of change of magnetic circuit geometry or of electromagnetic field configuration.The procedure is characterized by larger sensing ranges (see Figure 1.3) and lower energy consumption.

As previously mentioned, the system is self-generating and provides a measurand conversion by direct energy conversion. It can be used as stand-alone or as ISE of an IS (see Figure 1.2).

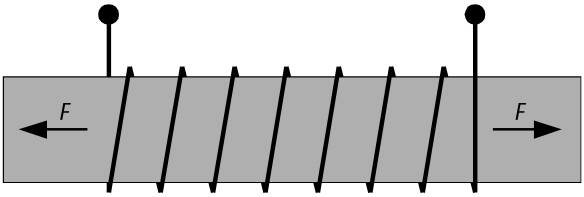


**Figure 3.2** Simple descriptive model of an magnetoelastic SE consisting of a coil wound onto a rod core of magnetostrictive material.

A first version, which is based on geometry influence, is shown in Figure 3.3. The target has a rigid mechanical linkage with a coil and moves this coil in a DC magnetic field that is provided by a permanent magnet.

The resulting induced voltage $v$ in the coil is generally expressed by the Faraday's law of electromagnetic induction:

$$v = -K\frac{\mathrm{d}\Phi}{\mathrm{d}t} \tag{3.5}$$

where $K$ is a physical constant determined by system properties and $\mathrm{d}\Phi/\mathrm{d}t$ represents time variation of the magnetic flux closed by the coil winding.

This voltage can be used as a stand-alone output signal or as a primary quantity in an IS (see Figure 1.2). Characteristic parameters are small elongations and low cutoff frequency.

An opposite functionality, namely the mechanical driving of a permanent magnet relative to a fixed coil is also met in practice. Both possibilities provide contactless sensing of any target (see Section 1.2.1).

For the second version of the electrodynamic systems detecting the change of the electromagnetic field configuration, the target becomes cooperative. It is made from ferromagnetic material and moves in the sensing lobe of an IS whose SE consists of fixed coil with permanent magnet core. The target movement causes periodical changes of the permanent magnet field, and induces the voltage $v$ (see (3.5)). The main advantage of this version is its contactless actuation character (see Section 1.2.1).

Technical publications, that is [6, 7], mention a particular realization called a system with variable reluctance to be included in this second class. An extension of the system using a second excitation coil make possible the AC contactless actuation of a target, which is again made from ferromagnetic material (see Figure 3.4). The advantage over the above-mentioned DC realization is the absence of a permanent magnet; as is well-known its DC magnet field attracts small iron parts, which cover the sensing face and alter the IS features.

The primary coil is supplied with AC current, generates an AC magnetic field, and induces a voltage $v$ in the secondary coil, which is dependent on the system geometry (see (3.5)), thus on the distance $s$. A second advantage of the AC operation mode is the possibility to detect dynamic target behaviors but also static situations.

The arrangement is equivalent to a transformer circuit having a variable low coupling factor between windings. The air space between target and IS active face

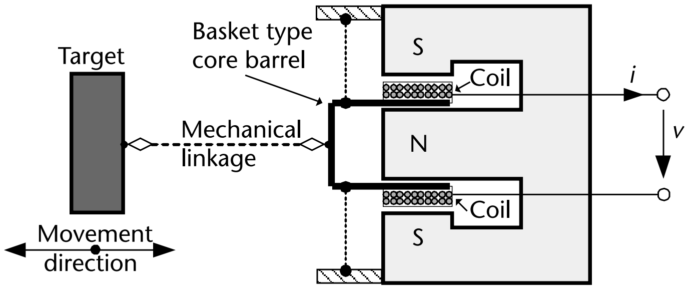


**Figure 3.3** Schematic diagram of an electrodynamic converter (longitudinal section view).

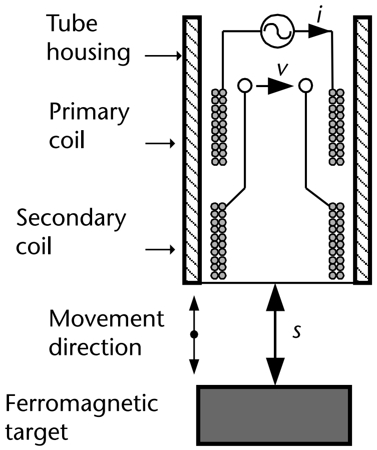


**Figure 3.4** Schematic representation of the system with variable reluctance (transversal section view).

(in front of the secondary coil) plays the role of a transformer gap and influences the magnetic circuit reluctance. Finally, there results a secondary voltage, which is dependent of the distance *s*. This dependence is unfortunately nonlinear but can be linearized using two inductive systems placed on both target sides. For this PP application, the primary coils are connected in series or parallel, while the secondary coils are ganged in the opposite direction.

As the introduction has already stated, one target of this book is to provide actual technical information about ISs using traditional analytical methods and descriptions but also the modern modeling and analyzing procedures offered by computer-aided design (CAD) techniques. We are starting here with the parallel description of the system in Figure 3.4 using the software Maxwell for computer-assisted simulation of electromagnetic fields [8].

For a moment, we are restricting the presentation to the final results of a first simulation project. The basics, processing, and evaluation of field simulation results will be explicitly presented in Section 4.3.

The system in Figure 3.4 is rotationally symmetric. In this case, to save time and money it is reasonable to consider a two-dimensional (2D) simulation model, which shows the right half of this system relative to its symmetry axis (see the representation in the right window in Figure 3.5). The simulation software considers a full rotation of the 2D model around the symmetry axis.

As it results from the project tree (left window in Figure 3.5), the model exemplifies a system with variable reluctance having the following specification:

- Two coils, each consisting of 144 turns from an enamel-insulated copper wire with a 0.2-mm diameter and disposed in 18 layers. Copper material parameters: relative permeability $\mu_r$ = 0.999991 and specific conductivity $\sigma$ = 5.8E7 S/m;
- Metal brass protective tube housing with $\mu_r$ = 1 and $\sigma$ = 1.5E7 S/m;
- Round plate target; material NiFe with $\mu_r$ = 4000 and $\sigma$ = 1.03E7 S/m;

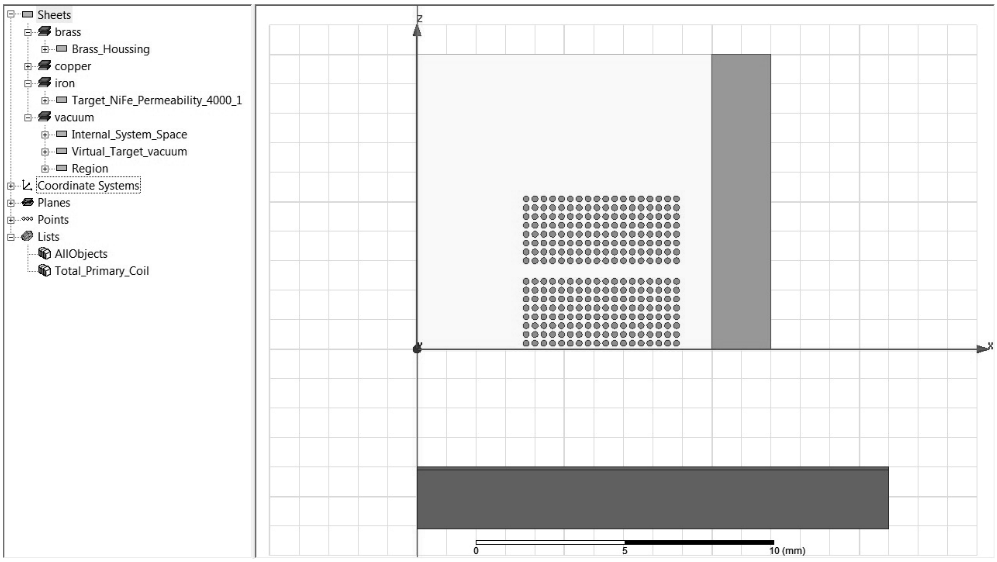

Sheets
brass
Brass_Houssing
copper
iron
Target_NiFe_Permeability_4000_1
vacuum
Internal_System_Space
Virtual_Target_vacuum
Region
Coordinate Systems
Planes
Points
Lists
AllObjects
Total_Primary_Coil
z
x
0
5
10 (mm)

**Figure 3.5** Maxwell field simulator 2D model of the system in Figure 3.4 (right half). For system dimensions, see the scale in the right window.

- Supply current conditions of the primary coil (upper position in Figure 3.5): 10 mA/10 kHz.

Figure 3.6 shows an example of the simulation results: the field line traces for a distance $s = 4$ mm between sensing face (on a level with $x$-axis) and target. The colors of the field lines mirror the magnetic flux value. The affinity between magnetic field lines and the ferromagnetic target is in evidence.

The display of the field line is the first step in the evaluation of the results provided by the simulation software. Its postprocessor delivers real values of the electromagnetic quantities (magnetic and electric fields, flux, flux density, etc.) in every point, on every surface or in every part, which belongs to the simulated region.

The author developed and published in team a method to achieve the most relevant results, namely the values of the equivalent electrical circuit parameters (e.g., impedance $\underline{Z}$ with its components resistance $R$ and reactance $X$ in (3.1), and inductance $L$ in (3.3), etc.). These features are discussed in Section 4.3 in greater detail.

### 3.3.3 Electromagnetic Systems with Closed Magnetic Loop

The inductive sensors with closed magnetic loop (IS_CML) define an inductive sensor class with a large occurrence. The sensor consists of a coil having a magnetic core with almost closed field lines.

Their theory of operation is based on the self-inductance variation due to the change of the system geometry (figurative variable $G$) in the general expression:

$$L = \text{function}(n, \mu, G) \tag{3.6}$$

where $n$ is the number of turns of the coil, $\mu$ is the core permeability and $G$ symbolizes the system geometry.

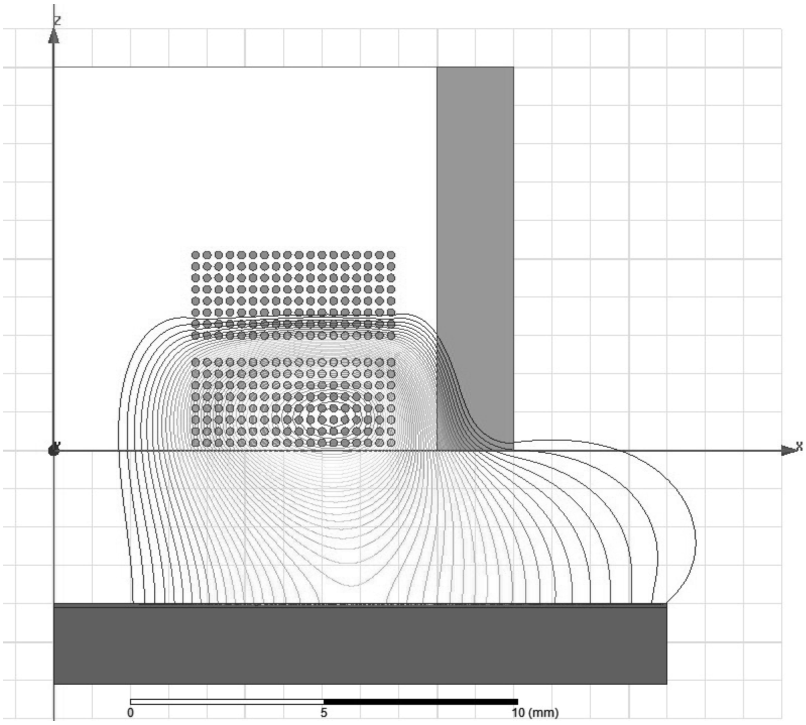


**Figure 3.6** Result of the Maxwell field simulator: snapshot of the field lines for an iron target at a distance $s = 4$ mm.

The device is passive and—as in Section 3.1 mentioned—provides a measurand evaluation by direct energy conversion. Therefore, it can be used as stand-alone, but also as an ISE of an IS (see Figure 1.2).

To describe the basic principle, a simple explanatory model is shown in Figure 3.7(a). A coil with $n$ turns is placed around a magnetic yoke core (cross section is rectangular, round, etc.), made from a ferromagnetic material. The air gap between this core and a mobile transversal armature (usually from the same material) represents the system measurand, namely the specific distance $s$.

The coil is usually one of the components of an LC oscillator and is supplied with the AC current $i$. It acts as a source of mmf, which drives the flux through the cores and air gaps. The flux lines preponderantly take course through medium with high magnetic permeability, namely, the core (see the dashed line for the typical field line trajectory in Figure 3.7(a)).

A good estimation of the behavior of this magnetic system can be achieved using the equivalent magnetic circuit, which is similar to an electrical circuit and relies on the analogy between electric and magnetic magnitudes (Figure 3.8).

Ohm's law is also valid for equivalent magnetic circuits. Thus, the magnetic flux $\Phi$ (see Figure 3.8) is expressed as:

$$\Phi = \frac{V_m}{R_m} \tag{3.7}$$

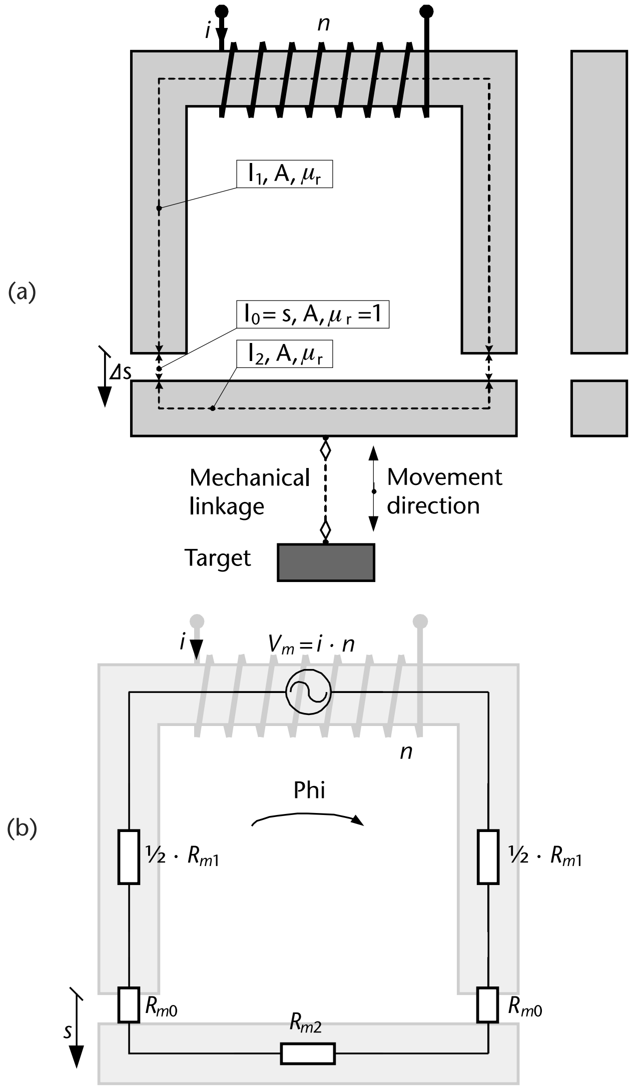

(a)
i
n
$l_1$, A, $\mu_r$
$l_0$= s, A, $\mu_r$ =1
$l_2$, A, $\mu_r$
Δs
Mechanical linkage
Movement direction
Target
(b)
i
$V_m = i \cdot n$
n
Phi
½ · $R_{m1}$
½ · $R_{m1}$
$R_{m0}$
$R_{m0}$
$R_{m2}$
s

**Figure 3.7** Schematic representation of the IS_CML: (a) Front view with the central flux path (dashed line) and side view, and (b) the ideal magnetic equivalent series circuit.

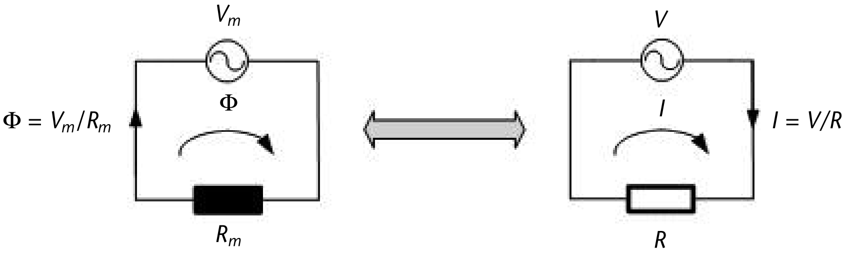

$V_m$
Φ
$\Phi = V_m/R_m$
$R_m$
V
I
$I = V/R$
R

**Figure 3.8** Analogy between the magnetic and the electric circuits.

where $V_m$ is the mmf and $R_m$ is the magnetic reluctance, generally defined by:

$$R_m = \int_{\substack{\text{magnetic}\\ \text{circuit}}} \frac{1}{\mu \cdot A} \mathrm{d}l \tag{3.8}$$

where $\mu$ and $A$ are the permeability ($\mu = \mu_0 \cdot \mu_r$) and the cross-sectional area of the magnetic circuit part. The integral is calculated for the entire flux path.

The resulting equivalent magnetic circuit of the IS_CML in Figure 3.7(a) is shown in Figure 3.7(b). It is important to mention that this view represents an ideal model with the following particularities:

- There are no magnetic losses.
- The system is linear, that is, the core relative permeability $\mu_r$ is constant.
- The stray flux does not exist. The stray flux is defined as the magnetic leakage flux, which is caused in real magnetic circuits by outward flux lines emanating from intermediate regions of this magnetic circuit, and which contributes to the reduction of the main flux in the magnetic circuit. A good illustration is shown in Figure 3.9. If the magnetic circuit is closed (zero air gap) there is theoretically no stray flux (see Figure 3.9(a)). As soon as an air gap is inserted (see Figure 3.9(b)), a lot of field lines exit the core forming the stray flux. Concluding, the stray flux is magnetic flux, which despite the efforts to guide it to where it has to go, does not go there.

The analytical modeling of the magnetic circuit in Figure 3.7(b) can be made based on the analogy shown in Figure 3.8, and using the Ohm's law for magnetic circuit (see (3.7)). Its equation reads as follows:

$$V_m = n \cdot i = \Phi \cdot \left(R_{m1} + R_{m2} + 2 \cdot R_{m0}\right) \tag{3.9}$$

where $R_{m1}$, $R_{m2}$, and $R_{m0}$ are the reluctances of the core, armature, and air gap, respectively (see Figure 3.7(b)).

Because the reluctance is generally inversely proportional to the magnetic permeability $\mu_r$ of the corresponding part, and since the core parts are made from materials having $\mu_r >> 1$, $R_{m1}$ and $R_{m2}$ are negligible and (3.9) can be reduced to:

$$n \cdot i = 2 \cdot \Phi \cdot R_{m0} \tag{3.10}$$

which emphasizes the fact that the magnetic flux along the system is inversely proportional to the air gap amount, namely to the sensing distance $s$:

$$\Phi = K_\Phi \cdot \frac{1}{s} \tag{3.11}$$

where $K_\Phi$ is a proportionality factor.

The transfer function of the IS_CML, namely the dependence of system inductance $L$ versus $s$, can be calculated using the ideal equivalent circuit in Figure 3.7(b) and expressing the total series reluctance $R_m$:

$$R_m = \sum_j R_{mj} = \frac{1}{A}\sum_j \frac{l_j}{\mu_j} \tag{3.12}$$

where $j$ is the sum index ($j = 0$ to 2), $A$ is the cross-sectional area of the flux path (assumed to be uniform over the entire circuit) and the values $l_j$ and $\mu_j$ are shown in Figure 3.7(a).

Consequently, the system inductance $L$ under steady-state conditions and considering core parts from the same material ($\mu_1 = \mu_2 = \mu_r$) is:

$$L = \frac{n^2}{R_m} = \frac{\mu_0 \mu_r A n^2}{l_1 + l_2 + 2\mu_r s} \tag{3.13}$$

Usually, the relative permeability of the core is of the order of a few thousand and $l_1$, $l_2$ lies in the range of a few centimeters. If the system geometry fulfils the condition:

$$\frac{l_1 + l_2}{\mu_r} \ll s \tag{3.14}$$

the system works efficiently (is air gap variation-sensitive) and its transfer function can be approximated by:

$$L \approx \frac{\mu_0 A n^2}{2s} \tag{3.15}$$

In conclusion, the presence of the air gap causes a large increase in circuit reluctance and a corresponding decrease in the flux. Hence, the small variation in the air gap causes a measurable change in inductance, which is the reason for [7] to name the system as IS with variable air gap.

The system sensitivity (see Section 1.3.3) for small target travel $\Delta s$ results from mathematically deriving the expression of the inductance $L$ in (3.15) and replacing the result of the derivation by an approximation using the Maclaurin series:

$$S \triangleq \frac{\Delta L}{\Delta s} \approx -\frac{\mu_0 A n^2}{2s^2}\left(1 - \frac{\Delta s}{s} \pm \ldots\right) \tag{3.16}$$

The expression shows a nonnegligible nonlinearity, since the sensitivity has a multiple dependence on the travel $s$. Therefore, the sensor is implementable for travels in millimeter ranges and for 5% to 10% $s$-jitter regarding the reference position.

In addition, the flux dispersion in the air gap area has an important influence and leads to significant system performance alterations.

Due to the stray of magnetic flux, also named dispersion effect of magnetic field, the effective flux $\Phi_a$ in the air gap is lower than $\Phi$ in the magnetic parts:

$$\Phi_a = \Phi(1-\sigma) \tag{3.17}$$

where $\sigma$ is the dispersion coefficient (usual values: 0.25 to 0.6 [2]). In a first approximation, the dispersion flux is independent of the air gap magnitude.

Consequently, the inductance increases with an amount dependent on the $\sigma$:

$$L = \frac{\mu_0 A n^2}{2s} + \frac{\mu_0 A n^2}{2s} \cdot \frac{\sigma}{1-\sigma} \tag{3.18}$$

Assuming a small target movement $\Delta s$ influences only the main flux and not the dispersion flux, the resulting new inductance is:

$$L(s+\Delta s) = \frac{\mu_0 A n^2}{2s} \cdot \left(1 + \frac{\sigma}{1-\sigma}\right) + \frac{\mu_0 A n^2}{2s}\left(-\frac{\Delta s}{s} \pm \ldots\right) \tag{3.19}$$

and allows to calculate the relative sensitivity of the system:

$$\frac{L(s+\Delta s) - L(s)}{L(s)} = -\frac{\Delta s}{s} \cdot (1-\sigma) \tag{3.20}$$

To minimize the dispersion effect, [3] recommends working at distances that fulfil the limitation:

$$s \leq \sqrt{0.05 \cdot A} \tag{3.21}$$

where $A$ is the cross-sectional area of the core.

Beside the nonlinearity, the system in Figure 3.7 is characterized by the existence of a magnetic force, which acts on the mobile transversal armature and tries to minimize the reluctance by reducing the air gap. The internal energy balance between work and energy consumption leads to the expression [3]:

$$F = -\frac{i^2 \cdot L \cdot \mu_r}{l_1 + l_2 + 2\mu_r s} \tag{3.22}$$

where $i$ is the supply current. This quite large force severely limits the system applications.

Figure 3.9 shows two examples of the simulation results, namely the real aspect of the field line for the sensing distances $s = 0$ mm and $s = 2$ mm, respectively. The colors of the field lines (in a color representation) mirror the magnetic flux value and, at the same time, give an image about the field distribution in the core parts.

The dispersion of magnetic field is obviously visible in Figure 3.9. If the air gap has the lowest technically possible value and the magnetic circuit is almost closed, there is practically no observable field dispersion (Figure 3.9(a)). As the air gap increases, the dispersion effect starts to show and the field stray grows according to the armature position.

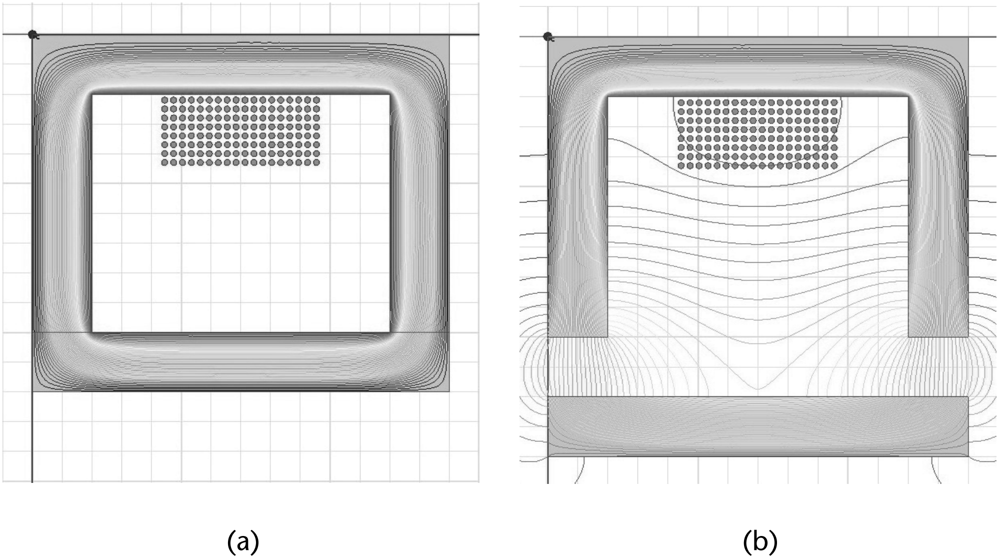


**Figure 3.9** Magnetic field lines (simulations results) for air gaps $s = 0$ mm and 2 mm, respectively (see Figure 3.7).

In conclusion, the reality could be very different from the idealized analytical model. The analytical system evaluation has its limits and the only possibility to overcome this obstacle is the using of the computer-aided field simulation tools.

A first improvement for better futures consists of a PP sensor version (see Figure 3.10). The system has a symmetrical topology, consists of two identical coils and

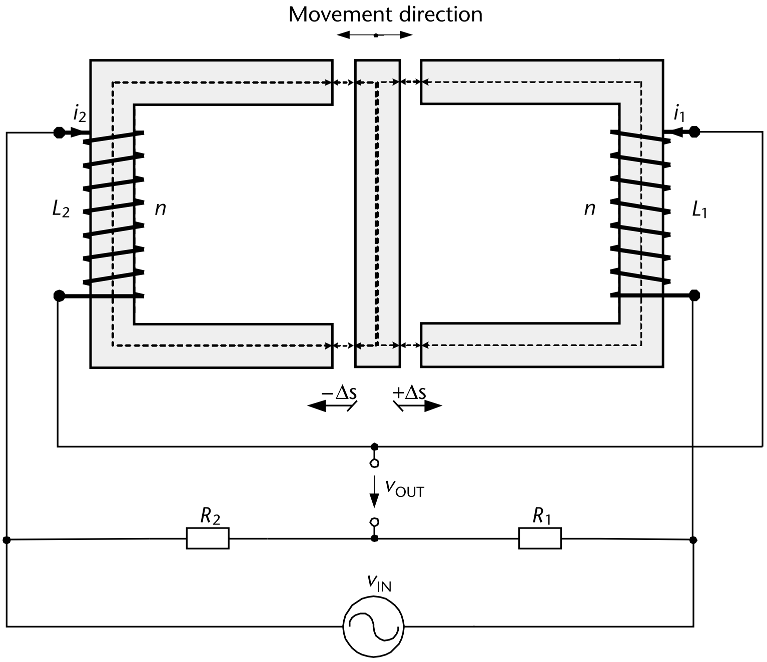


**Figure 3.10** Differential IS_CML (front view with the central flux paths (dashed lines)).

cores placed face-to-face, separated by a fixed distance. The mobile armature moves in the air space between cores in response to mechanical target and contributes to the reduction of the parasitic coupling between cores.

To evaluate the travel and the corresponding inductance changes, the coils can be incorporated into an AC deflection bridge, which gives a linear output for small movements.

The method provides the doubling of the sensitivity and the improvement of the linearity due to the elimination of the odd terms in the approximation series (see (3.16)). In fact, for the same small target movement $\Delta s$ (like in (3.19)) and for any parasitic coupling between coils, it results a difference $\Delta L$ between inductivities:

$$\Delta L = L(s + \Delta s) - L(s - \Delta s) = \frac{\mu_0 A n^2}{s} \cdot \left[ -\frac{\Delta s}{s} - \left( \frac{\Delta s}{s} \right)^3 - \left( \frac{\Delta s}{s} \right)^5 - \ldots \right] \tag{3.23}$$

The parasitic effect of the dispersions is eliminated and the linearity range is practically doubled. The system responds to both static and dynamic measurands. In addition, the force on armature and indirectly on target is substantially lower.

A second improvement valid for both versions described above consists of an E-shaped core instead of the U-shaped core [2]. The coil is placed on the central leg of the core. The flux lines symmetrically take course through both lateral core arms.

The analysis and modeling above are valid for low frequencies of the supply current (<50 kHz) and thereby have minor losses in the coils. For applications at higher frequencies, these losses can no longer be neglected. The target movement and the corresponding variation of the air gap $\Delta s$ changes both components of the coil impedance $\underline{Z}$ (compare with (3.1)):

$$\underline{Z}(s + \Delta s) = R_S(s + \Delta s) + j\omega \cdot L(s + \Delta s) \tag{3.24}$$

where $R_S$ and $L$ are the coil parameters series loss resistance and inductance, respectively, and $\omega$ is the angular frequency, which corresponds to the working frequency $f$:

$$\omega = 2\pi f \tag{3.25}$$

The series loss resistance $R_S$ basically consists of four components (for details, see Section 5.3.3), which are composed of the following effects:

1. Ohmic losses in the coil wire;
2. Eddy current losses;
3. Magnetic hysteresis losses;
4. Magnetic aftereffect (magnetic recovery).

For information only, there is a second geometrical possibility to change the inductance $L$ (see (3.15)) modifying the cross-sectional area $A$ of the flux path by a translation of the armature in parallel to the yoke core. The method provides larger ranges for the measured travel and has a better linearity due to the proportional dependence between $L$ and $A$. However, this method belongs to history now.

As a closing-off conclusion to Section 3.3.3, it could be mentioned that the IS_CML are suitable only for small to medium target travels (<10 mm) and have an acceptable sensitivity for such implementations. This type of IS particularly finds application in acceleration measurements. However, it can be designed to be suitable for travels or speeds sensing.

### 3.3.4 Electromagnetic Systems with Open Magnetic Loop

The inductive sensors with open magnetic loops (IS_OML) are behavior-related near the sensors with closed magnetic loops (see Section 3.3.3). The major difference results from the course of magnetic lines of force, which by now occur preponderantly in air and not in magnetic mediums.

Also named a system with mobile plunger (core), the device contains a solenoid and a mobile cylindrical-shaped ferromagnetic core, which moves inside the coil along its symmetry axis (see Figure 3.11). The collinear displacement of the core, which is driven by an external target having a rigid mechanical linkage to the core, causes the system coil inductance to change.

The system performs in this way a contactless sensing (see Section 1.2.1).

The device is passive and—as Section 3.1 mentioned—provides a measurand evaluation by direct energy conversion. Therefore, it can be used as stand-alone but also as an ISE of an IS (see Figure 1.2).

An analytical evaluation of the functional model (see Figure 3.11) with sufficient accuracy can be made using the following simplifying hypotheses:

- The magnetic field inside the coil has a uniform distribution and its strength is preponderant compared to the external field strength around the coil.
- Due to the relative low working frequencies, the eddy currents in the core windings can be neglected.
- For the same reason, the skin effect is also negligible. Skin effect is defined as the tendency of an AC current to become distributed within a conductor so that the current density is larger near the surface of the conductor, and decreases with greater depths in the conductor. The electric current flows mainly at the skin of the conductor, between the outer surface and a level called the skin depth (see (3.30)).
- The flux is conservative and, therefore, Ampère's circuital law:

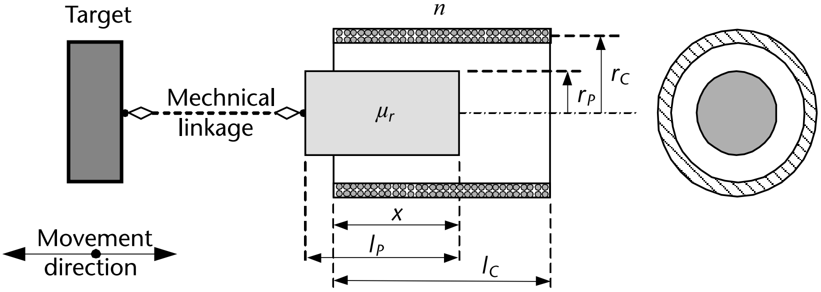


**Figure 3.11** Schematic representation of the IS_OML (longitudinal and cross section, respectively).

$$n \cdot i = \oint |H| \cdot \mathrm{d}l \tag{3.26}$$

with $n$ = number of the coil turns and $i$ = current flowing through the coil, is fulfilled independently of the closed integration path.

By this means, a quite laborious computation leads to the following expression of field strength inside the coil $H_{iC}$ [2]:

$$H_{iC} = \frac{n \cdot i}{\left[ l_C - x + \dfrac{x}{1 + \left(r_P / r_C\right)^2 \cdot \left(\mu_r - 1\right)} \right]} \tag{3.27}$$

where $l_C$ and $r_C$ are the coil length and middle radius, respectively. Values $r_P$ and $\mu_r$ are the plumber characteristics radius and relative permeability, respectively.

Inserting this expression into (3.3), it results the final formula for the coil inductance $L$ [2]:

$$L = K \frac{\mu_0 \cdot n^2 \cdot \pi r_C^2}{l_C} \cdot \frac{1}{1 - \dfrac{x}{l_C} \left[ 1 - \dfrac{1}{1 + \left(r_P / r_C\right)^2 \cdot \left(\mu_r - 1\right)} \right]} \tag{3.28}$$

Usually, the plumber has a high relative permeability ($\mu_r > 10\text{E}3$) and the inductance expression simplifies:

$$L \approx K \frac{\mu_0 \cdot n^2 \cdot \pi r_C^2}{l_C} \cdot \frac{1}{1 - \dfrac{x}{l_C}} \tag{3.29}$$

where the factor $K$ ($K < 1$) implies the errors due to the aforementioned hypotheses.

The arrangement shown in Figure 3.11 becomes a basic IS_OML if the immersion depth $x$ is allowed to vary (independent variable).

To continue doing a parallel system description (the procedure started in Section 3.3.2) of analytic versus computer-assisted, Figure 3.12 shows the used vertical model to simulate the IS_OML in Figure 3.11. The previous considerations to Figure 3.9 remain valid.

As the project tree (left window in Figure 3.12) summarizes, the model consists of:

- A system coil having 288 turns from enamel-insulated copper wire with a 0.2-mm diameter and disposed in 8 layers. Copper material parameters of relative permeability $\mu_r = 0.999991$ and specific conductivity $\sigma = 5.8\text{E}7$ S/m.
- Cylindrical plunger (core) from NiFe with $\mu_r = 4000$ and $\sigma = 1.03\text{E}7$ S/m.
- Supply current conditions of the coil: 10 mA/10 kHz.

Figure 3.13 shows an example of the simulation results: the field line traces for three significant positions of the core in relation to the coil, namely upper, middle,

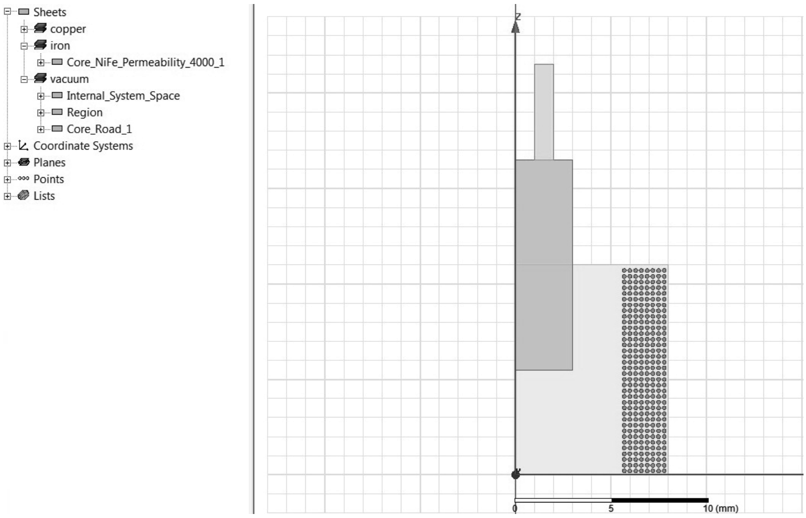


**Figure 3.12** Maxwell field simulator 2D model (vertical position) of the system in Figure 3.11 (right half). For system dimensions, see the scale in the right window.

and lower positions. The colors of the field lines (in a color representation) mirror the magnetic flux value. The affinity between magnetic field lines and the ferromagnetic core as well as the skin effect in the core are in evidence.

As already mentioned, the field simulation offers a large variety of results. Figure 3.14, for example, shows the magnetic field vector **H** inside the sensor at the middle core position. This vector is represented by oriented arrows, which are tangential to the field lines. The vector-magnitude is given by the length of the arrows.

At high frequencies, the skin effect makes an appearance. The field concentrates in a thin region under the plunger surface and the plunger is physically equivalent to a thin cylinder with the same relative permeability $\mu_r$, approximately. The thickness of this region, called skin depth, is given by the classic formula:

$$\delta = \sqrt{\frac{2}{\omega \cdot \mu \cdot \sigma}} \tag{3.30}$$

where $\omega$ is the angular frequency (see (3.25)), $\mu$ and $\sigma$ are the permeability and the specific conductivity of the plunger material, respectively.

To calculate the inductance $L$ of this resulting system with thin cylindrical plunger, [3] proposes to divide it in three components (see Figure 3.15):

- A first coil with a cylinder-shaped core, with the length equal to the plunger immersion depth $x$ and having a resulting turns number $n_1$;
- A second coil without a core and having the remainder number of turns $n_2$ ($n_2 = n - n_1$);
- A nonsignificant remainder core segment. Reference [3] proposes to neglect the core segment without coil and to connect in series the two coils.

For this equivalent topology, [3] releases the following equation for the total system inductance $L$:

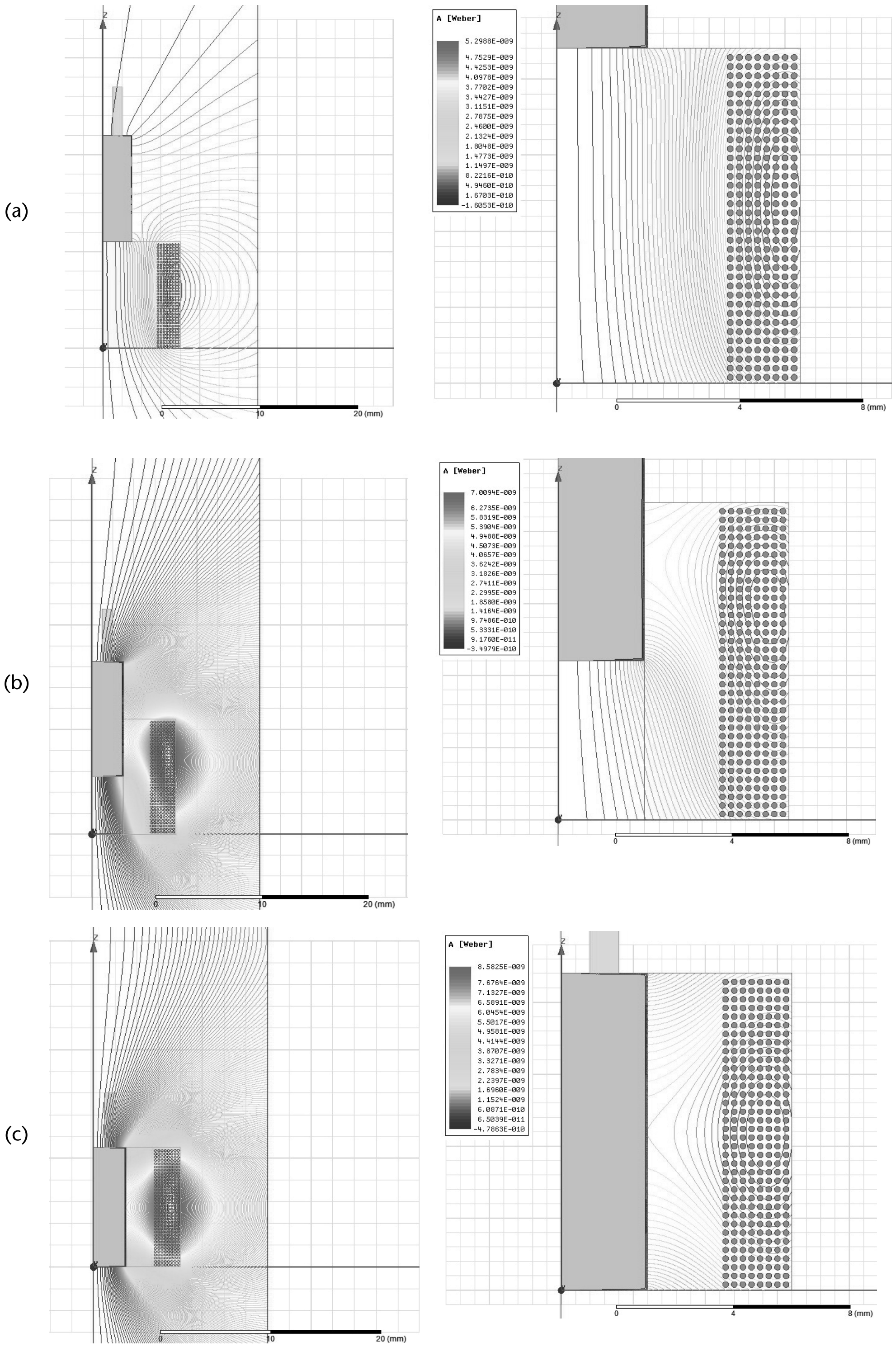


**Figure 3.13** Result of the Maxwell field simulator: full-range image (left column) and magnified snapshots of the field lines (right column) for three significant positions of the plunger in relation to the coil: (a) Upper plunger position (this position corresponds to the minimal distance (LwD) in Figure 2.6), (b) middle plunger position, and (c) lower plunger position (this position corresponds to the maximal distance (UpD) in Figure 2.6). The three spectral scales cover the same range.

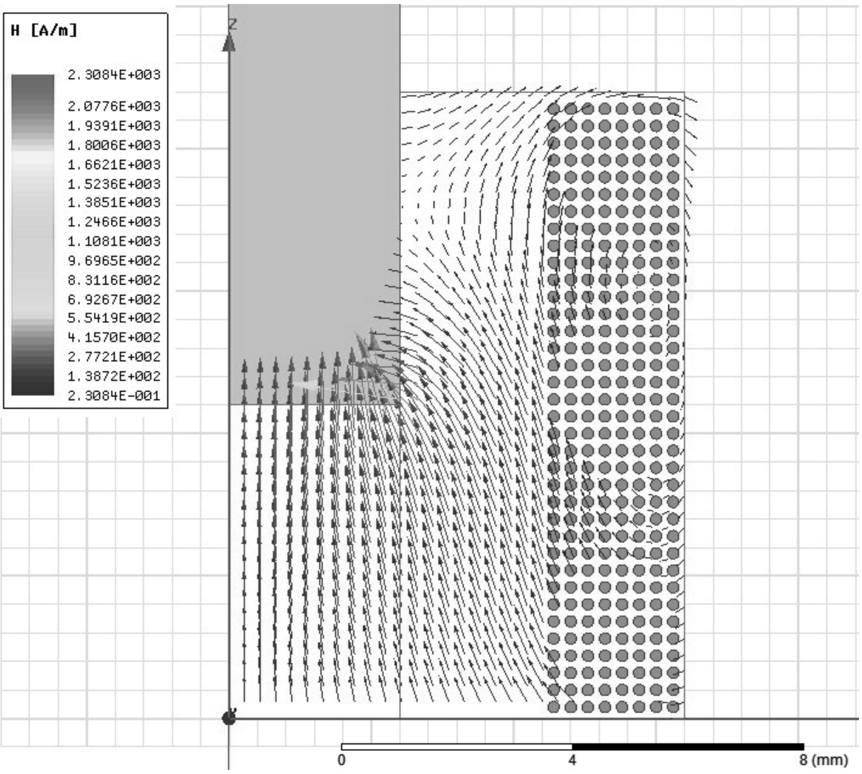


**Figure 3.14** Result of the Maxwell field simulator: Magnetic field vector **H** inside the sensor and for the middle plunger position.

$$L = \frac{4\pi^2 \cdot n^2 \cdot r_C^2}{l_C} \cdot \left[ 1 + \left( \frac{r_P}{r_C} \right)^2 \cdot \left( \frac{\delta \cdot \mu_r}{r_P} - 1 \right) \cdot \frac{x}{l_C} \right] \cdot 10^9 \text{ [henry]} \tag{3.31}$$

where the measuring unit for all geometrical values is the cm. The variable $x$ is plunger immersion depth. For the remaining notations, see (3.27).

For high permeability plunger materials ($\mu_r > 10\text{E}3$), the relation (3.31) becomes:

$$L \approx \frac{4\pi^2 \cdot n^2 \cdot r_C^2}{l_C^2} \cdot \left( \frac{r_P}{r_C} \right)^2 \cdot \left( \frac{\delta \cdot \mu_r}{r_P} \right) \cdot x \cdot 10^9 \quad \text{[henry]} \tag{3.32}$$

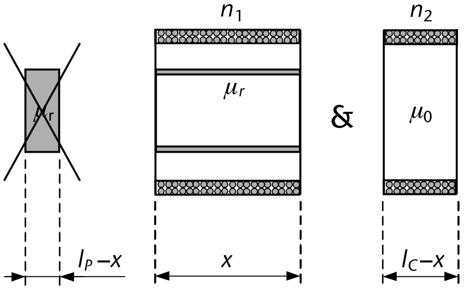


**Figure 3.15** Approximation of the core coil in Figure 3.11 by a series of two coils.

and evidences a direct proportionality between inductance and target travel as well as the geometrical possibilities to improve the system sensitivity:

$$S \triangleq \frac{\Delta L}{\Delta x} \approx \frac{4\pi^2 \cdot n^2 \cdot r_C^2}{l_C^2} \cdot \left(\frac{r_P}{r_C}\right)^2 \cdot \left(\frac{\delta \cdot \mu_r}{r_P}\right) \cdot 10^9 \ \frac{\text{henry}}{\text{cm}} \tag{3.33}$$

The single-coil version (see Figure 3.11) can be improved using two identical coils in a differential arrangement. The system is also called a linear variable inductor. A typical arrangement consists of two collinear, identical center tapped coils and a ferromagnetic plunger with the same length (see Figure 3.16). The core translates left and right relative to the system median line and provides the mechanical input. The two coils are placed to form two arms of a bridge circuit with two equally balancing resistors, $R_1$ and $R_2$. The bridge performs the supply of the coils and the output evaluation.

In the middle null position, the output voltage is equal to zero. Movements cause the output voltage to vary and to change the sign dependent on the movement direction.

The bridge is excited with AC of 5 to 25V with a frequency between 50 Hz and 5 kHz. The resistors should have the same value as the coil impedances in null position (100 to 1,000Ω) and the output load must be at least 10 times the resistors value.

Specific for differential procedures, the system has a better linearity (typical 1% full-scale linearity error) and a lower temperature dependence.

The IS_OML can be used for travels in a wide range, from millimeter-range up to the order of a few hundreds of centimeters. For large travels, the asymmetrical version is preferred in order to keep reasonable system dimensions.

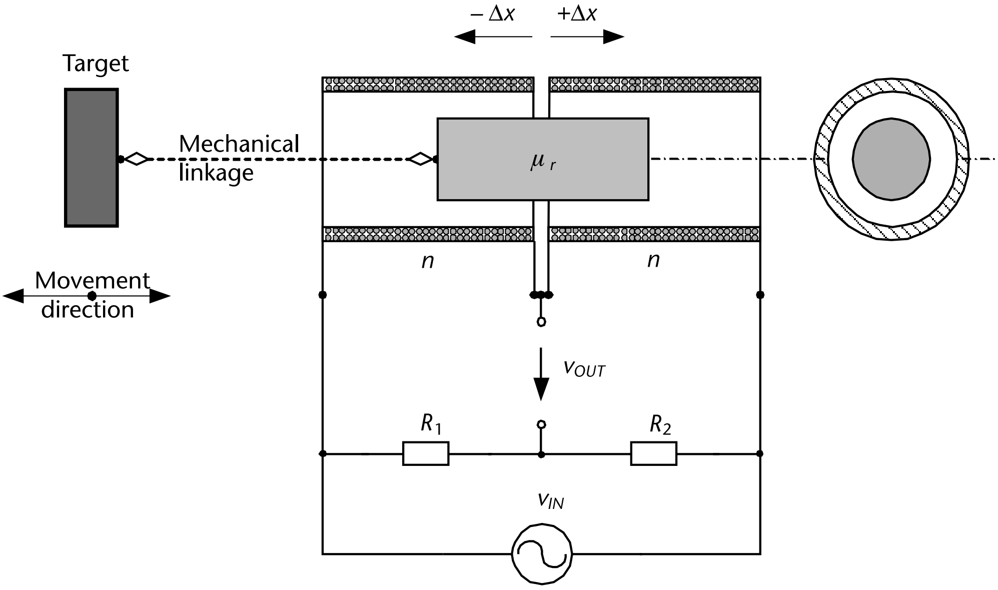


**Figure 3.16** Symmetrical arrangement (longitudinal and cross section, respectively) of the sensors with open magnetic loop shown in Figure 3.11 and the AC deflection bridge to evaluate the coil inductivities difference.

The sensor is available in various shapes and sizes. Rotary types are available as well. Over a range of 90° rotations, the sensitivity can be up to 100 mV per degree of rotation.

Compared with the IS_CML, the IS_OML has a lower quality factor, due to the open magnetic circuit. Despite that, they are preferred in numerous industrial applications due to their robustness, high resolution and repeatability, and resilience against core rotation. The cylindrical configuration can be better encased and protected against harsh environments.

### 3.3.5 Variable Differential Transformers

A short view of the Figure 3.17 and the comparison with representation in Figure 3.16 gives the impression that the variable differential transformer (VDT) is very similar to IS_OML. In fact, they are of quite a different nature. As the IS_OML is a variable-inductance system and its output is the inductance, the VDT is a variable-coupling device having a voltage output and being based on the transformer technology.

Transformers work by exciting the primary winding with an AC voltage and inducing a voltage in the secondary winding by subjecting it to the changing magnetic field set up by the primary. They are inductive by nature, consisting of wound coils.

By varying the amount of coupling from the primary (excited) winding to the secondary (coupled) winding with respect to either linear or rotary displacement, an analog signal that represents the displacement can be generated. The coupling variation is accomplished by moving either one of the windings or a core element that provides a flux path between the two windings.

A popular form of an electromagnetic position/displacement sensor is the VDT with its versions: linear (LVDT) [9] and rotary (RVDT), respectively.

A preliminary presentation of the VDTs was inserted here to ensure a consistent content of the Section 3.3, in other words, to have all IS types in a body.

The schematic representation of the classical LVDT in Figure 3.17 evidences the single primary winding positioned between two secondary windings wound on a tubular insulating former. These are made as identical as possible by having equal sizes, shapes, and numbers of turns. The cylindrical ferromagnetic core having the already known rigid linkage with the target performs a longitudinal movement inside the former and along the system symmetry axis. The core is prevented from parasitic rotation by a longitudinal narrow slot.

The primary winding is excited by AC voltage (frequency up to 20 kHz) and the secondary windings are connected in series opposition.

As the core inside the former moves, the magnetic paths between primary and secondary windings change, thus giving secondary outputs proportional to the movement. In the null core position, the differential output voltage $v_{\text{out}}$ is zero. Supposing the core moves to the left, from this null, more magnetic flux links with the left-hand coil and the voltage induced in the left-hand coil is then larger than the induced voltage on the right-hand coil. It results an output voltage having a larger value than at the null position and having a phase, which indicates the left-hand movement. Similarly, the movement in opposite direction reverses the effect and the output voltage is now in phase with the induced voltage on the right-hand coil.

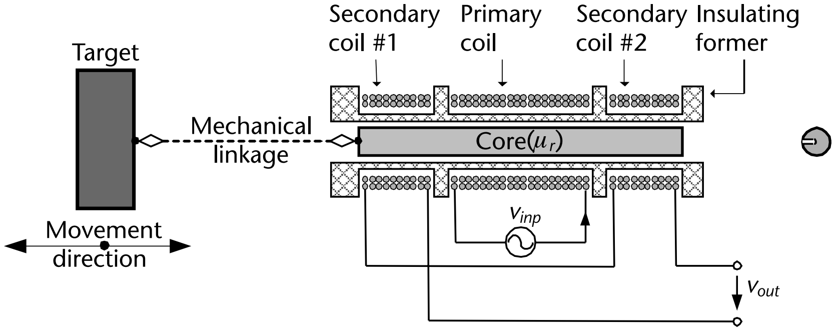

Target
Secondary coil #1
Primary coil
Secondary coil #2
Insulating former
Mechanical linkage
Core($\mu_r$)
$V_{inp}$
Movement direction
$V_{out}$

**Figure 3.17** Schematic representation of the LVDT (system longitudinal section and core cross section, respectively).

To continue in this section, the chosen methodology of physical versus computer-assisted description, Figure 3.18 shows the used functional model in a vertical position to simulate the LVDT from Figure 3.17. The previous considerations remain valid.

The project tree (left window in Figure 3.18) describes the relevant model components and simulation conditions:

- Three system coils having each 288 turns from enamel-insulated copper wire with a 0.2-mm diameter and disposed in 8 layers. Copper material parameters: relative permeability $\mu_r = 0.999991$ and specific conductivity $\sigma = 5.8E7$ S/m;
- Cylindrical housing from Al with $\mu_r = 1.00021$ and $\sigma = 3.6E7$ S/m;
- Cylindrical plunger (core) from NiFe with $\mu_r = 4000$ and $\sigma = 1.03E7$ S/m;
- Parameters of the supply current of the primary coil: 25 mA/2.5 kHz (very closed to a real product).

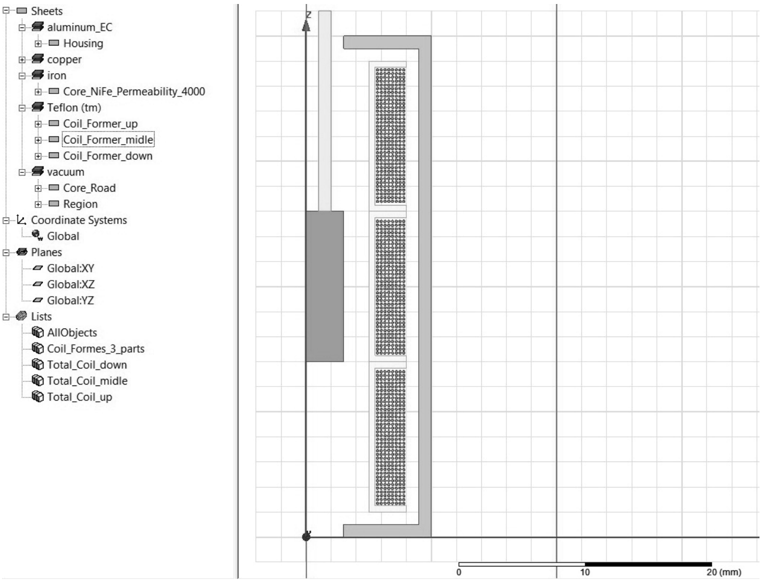

Sheets
aluminum_EC
Housing
copper
iron
Core_NiFe_Permeability_4000
Teflon (tm)
Coil_Former_up
Coil_Former_midle
Coil_Former_down
vacuum
Core_Road
Region
Coordinate Systems
Global
Planes
Global:XY
Global:XZ
Global:YZ
Lists
AllObjects
Coil_Formes_3_parts
Total_Coil_down
Total_Coil_midle
Total_Coil_up
z
0
10
20 (mm)

**Figure 3.18** Maxwell field simulator 2D model (vertical position) of the encapsulated LVDT in Figure 3.17 (right half). For system dimensions, see the scale in the right window.

Figure 3.19 shows an example of the simulation results: the field line traces corresponding to two significant positions of the core in relation to the coils, namely middle and lower positions. The colors of the field lines (in a color representation) mirror the magnetic flux value. The affinity between magnetic field lines and the ferromagnetic core as far as the skin effect are in evidence.

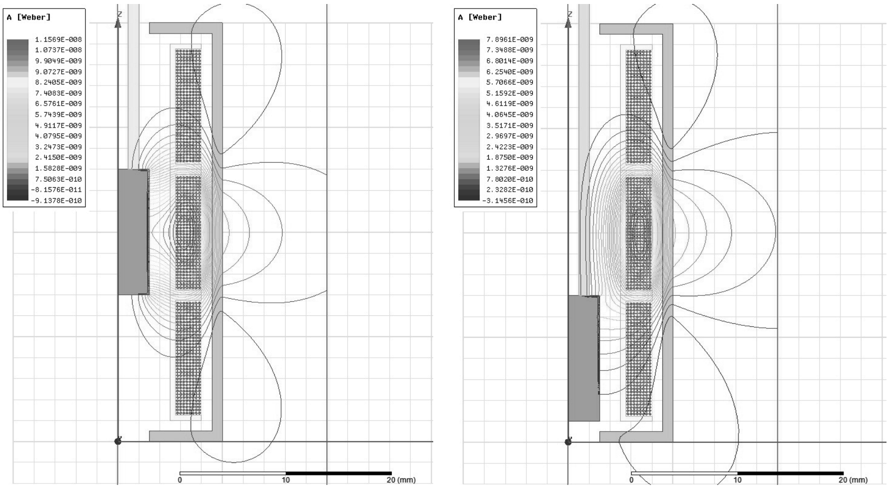
The RVDT consists of an E-shaped core and a slim ferromagnetic armature. The primary winding is wound on the center core leg and the secondary windings are placed on the outer legs of the core. The supply and the wiring are identical to the LVDT.

The RVDT operates in a similar manner as the LVDT. The armature is rotated by an externally applied force about a pivot point above the coil leg center in a range of −45° to +45°. When the armature is displaced from its reference position, the reluctances of the magnetic circuit through one secondary coil decreases and the other coil increases. Consequently, the induced voltages in the secondary windings, which are equal in the armature reference position, become different in magnitude and phase. Finally, a differential output voltage results that similar to $v_{\text{out}}$ (see Figure 3.17) is not equal to zero and has a phase, which indicates the rotation direction.

### 3.3.6 Systems Based on the Eddy Currents Evaluation

The common characteristic of the previous systems (see Sections 3.3.1 through 3.3.5) is the influence of the target movement on the inductance of a winding. At the specific low working frequencies, the influence on the winding losses is reduced and can be neglected.

In opposition to these devices, the inductive sensors based on eddy currents (IS_EC) work at high frequencies and make primarily use of the losses evaluation. These losses are generally dependent on the distance target to coil and are measured by the IS_EC.


A [Weber]
1.1569E-008
1.0737E-008
9.9049E-009
9.0727E-009
8.2405E-009
7.4083E-009
6.5761E-009
5.7439E-009
4.9117E-009
4.0795E-009
3.2473E-009
2.4150E-009
1.5828E-009
7.5063E-010
-8.1576E-011
-9.1378E-010
Z
0
10
20 (mm)
A [Weber]
7.8961E-009
7.3488E-009
6.8014E-009
6.2540E-009
5.7066E-009
5.1592E-009
4.6119E-009
4.0645E-009
3.5171E-009
2.9697E-009
2.4223E-009
1.8750E-009
1.3276E-009
7.8020E-010
2.3282E-010
-3.1456E-010
Z
0
10
20 (mm)

(a) (b)

**Figure 3.19** Result of the Maxwell field simulator: snapshots of the field lines for two significant positions of the core, (a) middle core position, and (b) lower core position.

This sensor type is certainly one of the largest used sensors in industrial applications. The annual volume of produced and supplied devices is of the order of a several millions. At least ten global players are involved in this business (see Section 3.4.3).

The IPS, whose characteristics were described in detail in Chapter 2, is the dominant IS_EC, by far. It is a market leader and provides the best and safe contactless actuation for theoretically whole plenty of conductive, ferromagnetic or nonferromagnetic materials. The commercial IPSs have the entire spectrum of output types (see Section 1.3.7) and are used either for position or for displacement sensing [10]. This sensor family represents a cornerstone of the present book.

The classical SE implementation of the IS_EC consists of a sending/receiving coil, with or without a magnetic core, excited by HF AC currents (usually in MHz range), placed in front of the object (target) to be detected and generating an HF AC magnetic field. Topologies with ferrite pot core (Figure 3.20) are preferred because they are able to orientate and focus the magnetic field lines on the target and to provide a certain electromagnetic shielding from metal environment MW (see Figure 2.2).

When the target approaches the coil and enters in its sensing range, the generated magnetic flux passes into conductive target. It links with the target and produces eddy current within this. The eddy current density is significant at the target surface and exponentially drops in a thin target layer below the surface. After a depth $a$ equal to five skin depths (see (3.30)):

$$a = 5 \cdot \delta = 5 \cdot \sqrt{\frac{1}{\pi \cdot f \cdot \mu \cdot \sigma}} \tag{3.34}$$

where $f$ is the excitation frequency, $\mu$ and $\sigma$ are the permeability and the specific conductivity of the target material, respectively, this density is negligible. The total depth of this layer lies in submillimeter range; electromagnetic phenomena exclusively occur in this region.

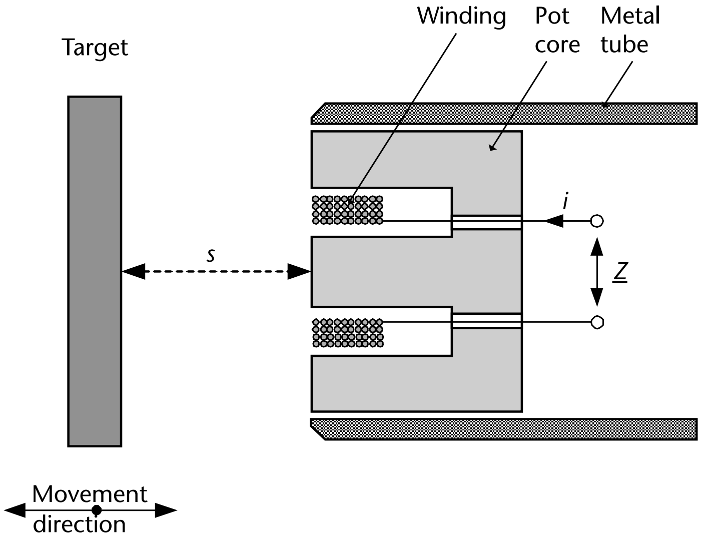


Figure 3.20 Schematic representation of the IS_EC (transversal section).

As the target comes closer to the coil, the eddy currents become stronger, causing a change in the coil impedance $\underline{Z}$:

$$\underline{Z}(s) = R_S(s) + j\omega \cdot L(s) \tag{3.35}$$

where the variable $s$ is the distance between target and coil, $R_S$ and $L$ are the coil parameters series loss resistance and inductance, and $\omega$ is the angular frequency, which corresponds to the working frequency $f(\omega = 2\pi f)$.

Both parameters $L$, but predominantly the series loss resistance $R_S$, are affected. As already mentioned in Section 3.3.3, the loss resistance $R_S$ is the sum of four components, which are composed of the already-mentioned effects and are independently influenced.

Figure 3.21 shows the equivalent electrical series circuit of the coil according to (3.35) and a very common method to supply the coil by means of an LC oscillator (see Chapter 8).

Making use again of the field simulator, Figure 3.22 shows the magnetic flux in a 2D spectral representation and keeps alive the assumption that a relevant flux spreading in target occurs inside a ring zone (where $\Phi$ = 4E-11 Vs typ).

The ring with the thickness $a$ and radii dependent on the coil dimensions and distance $s$ is equivalent with a single turn winding working in short-circuit conditions. The equivalent circuit in Figure 3.21 can be now followed up replacing the target (with its virtual ring) by this winding. The winding itself has an equivalent series circuit with the components $L_r$ and $R_r$.

It results in a new, fully electrical equivalent circuit, which corresponds to a transformer working in short-circuit conditions (see Figure 3.23). To simplify the evaluation, the secondary current $i_2$ is considered constant, although the current density exponentially drops within the ring depth. The secondary circuit parts indirectly reflect ferro-, dia-, or paramagnetic target material properties.

Finally, the eddy currents induced in the target generate a new magnetic field—secondary magnetic field—that is in phase opposition to the main primary magnetic field and has retroactive effects on the primary components $L$ and $R_S$.

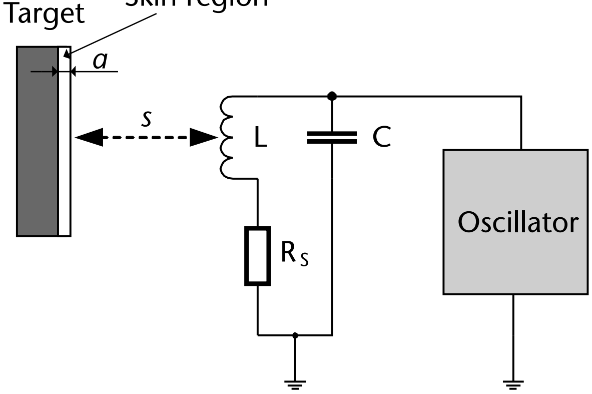


**Figure 3.21** Excitation of the SE of an IS_EC by a dipole oscillator (SE is represented by the series circuit $L$ and $R_S$).

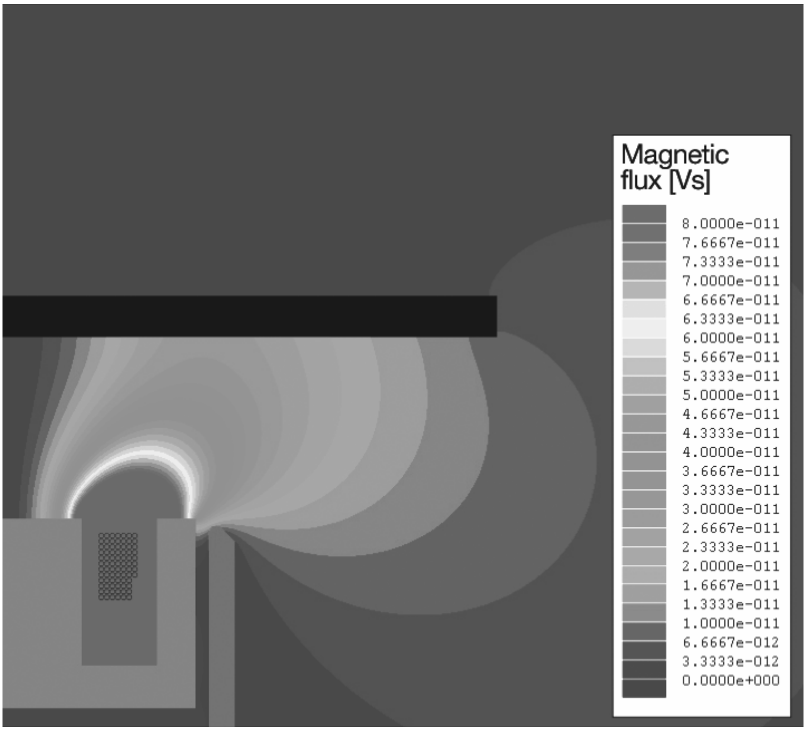


**Figure 3.22** Magnetic field of the rotation-symmetrical coil in pot core (right half) of an IS_EC in vertical position (the target diameter is larger than the coil diameter).

The mathematical analysis of the operation starts with the classical definition of the coupling factor $k$:

$$k = \frac{M}{\sqrt{L \cdot L_r}} \tag{3.36}$$

where $M$ is the mutual inductance.

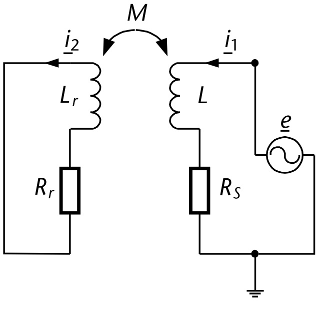


**Figure 3.23** Final equivalent electrical circuit of the arrangement containing the SE of an IS_EC (components $L$ and $R_S$) and target (represented by $L_r$ and $R_r$).

Solving the complex transformer equations; that is:

$$\begin{aligned}(R_S + j\omega L)\cdot \underline{i}_1 + jwM\cdot \underline{i}_2 &= \underline{e}\\ (R_r + j\omega L_r)\cdot \underline{i}_2 + jwM\cdot \underline{i}_1 &= 0\end{aligned} \tag{3.37}$$

the expression of the primary impedance results, which loads the excitation source $\underline{e}$:

$$\underline{Z} \overset{\text{def}}{=} \frac{\underline{e}}{\underline{i}_1} = R_S + \frac{\omega^2 M^2}{R_r^2 + \omega^2 L_r^2}\cdot R_r + j\omega\left(L - \frac{\omega^2 M^2}{R_r^2 + \omega^2 L_r^2}\cdot L_r\right) \tag{3.38}$$

In conclusion, due to the electromagnetic linkage with the target, the equivalent loss resistance results higher:

$$R_{S\_\mathrm{eq}} = R_S + \frac{\omega^2 M^2}{R_r^2 + \omega^2 L_r^2}\cdot R_r \tag{3.39}$$

and the equivalent inductance results lower:

$$L_{\mathrm{eq}} = L - \frac{\omega^2 M^2}{R_r^2 + \omega^2 L_r^2}\cdot L_r \tag{3.40}$$

The opposite variations in these directions describe behaviors, which are specific for the eddy current effects in usual conductive materials.

When the following inequality is fulfilled (high working frequencies):

$$R_r \ll \omega L_r \tag{3.41}$$

the equations (3.39) and (3.40) can be simplified and reorganized to highlight the coupling factor *k* and—this way—to underline the travel *s* dependence of the equivalent parameters:

$$R_{S\_eq}(s) = R_S + k^2\frac{L}{L_r}\cdot R_r \tag{3.42}$$

$$L_{\mathrm{eq}}(s) = L(1 - k^2) \tag{3.43}$$

where the coupling factor *k* increases when the distance *s* decreases.

This modeling does not explicitly evidence the magnetic properties of the target. In reality, these equations indirectly imply the target permeability $\mu$. The coupling factor *k* is dependent on the thickness *a* of the effective virtual ring and *a* is a function of $\mu$ (see (3.34)).

In addition, for coils with ferrite pot core (see Figure 3.20) and ferromagnetic targets situated at small distances, the specific eddy currents behavior is suppressed by a predominant effect specific to sensor systems with a closed magnetic loop (compare with Section 3.3.3).

To capture both components simultaneously (see (3.42) and (3.43)), it is advised to introduce the notion of quality factor. The particular expression for the quality factor of the coil $Q_L$ results returning to the primary impedance $\underline{Z}$ (see (3.38)):

$$Q_L \overset{\text{def}}{=} \frac{1}{D} = \frac{1}{\operatorname{arctg}\dfrac{\Im \underline{Z}}{\mathcal{R}\underline{Z}}} \approx \frac{\Im \underline{Z}}{\mathcal{R}\underline{Z}} = \frac{\omega \cdot L_{\text{eq}}(s)}{R_{S_{\text{eq}}}(s)} \tag{3.44}$$

where $D$ is the coil loss factor, classically defined by means of real and imaginary parts of the primary impedance $\underline{Z}$.

In rare cases, the SE of the IS_EC consists of two coils. The first one is similar to the above-described coil and plays the same role. The second is a balance coil placed back-to-back behind the active coil. It is not actuated by the target but it is influenced by temperature, making possible efficient temperature compensations at large distances.

Short summaries of the standard sizes of the IPSW based on IS_EC with the conventional features are already available in Sections 2.1.1.6 and 2.1.1.9.

### 3.3.7 Variable Transformers: Microsyn, Synchro, and Resolver

The devices called variable transformers (VT) (microsyn, synchro, resolver, etc.) have a special status. On one hand, they are considered a particular sensor family on the large inductive sensor field, which is a reason for us to mention and describe it. On the other hand, they can be deemed as a particular category of electrical machines.

These rotating machineries, which have a respectable age in various applications, show significant features for the rotation and angle measurements and can stand any competition, like modern optical rotary encoders or potentiometric angle transducers.

Due to their relevant advantages, several constructive versions are designed, fabricated and offered on the market. The present section describes the most important, popular three VT types.

#### 3.3.7.1 General VT Theory of Operation

The concept of VT basically describes a classical transformer whereby the relative position between primary and secondary windings is variable (rotation and/or translation). The movement of one winding, relative to the second winding, changes the coupling factor $k$ between them (see (3.36)). The $k$-value variation changes the mutual inductance $M$ and, when one winding is excited, the induced voltage in other winding varies dependently on $M$.

The schematic representation in Figure 3.24 shows a possibility to change the parallelism between the symmetry axes of the windings, rotating the secondary by the angle $\alpha$ (alpha).

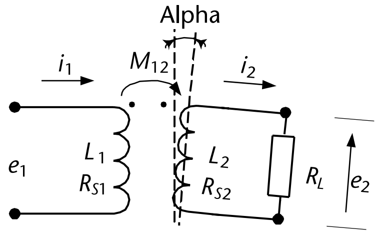


**Figure 3.24** Schematic winding diagram of a VT with variable angle $\alpha$ between windings ($L_1$, $R_{S1}$ are primary winding, $L_2$, $R_{S2}$ are secondary winding).

The mutual inductance $M_{12}$ is expressed as:

$$M_{12} = n_2 \frac{\Phi_2}{i_1} \tag{3.45}$$

where $n_2$ is the turns number of the secondary winding, $i_1$ the current in the primary winding and $\Phi_2$ the magnetic flux in the secondary.

A rotation of the secondary modifies the effective cross section and consequently changes the magnetic flux in a trigonometric manner:

$$\Phi_2 = \mu \frac{n_1 i_1}{l} A \cdot \cos\alpha \tag{3.46}$$

where $A$ and $l$ are the cross-sectional area and the length of the secondary, $n_1$ the turns number of primary, and $\mu$ the medium permeability. Writing (3.46) in (3.45) results an expression that shows the exclusive geometrical dependence of the system mutual induction:

$$M_{12} = n_1 n_2 \frac{\mu}{l} A \cdot \cos\alpha = m \cdot \cos\alpha \tag{3.47}$$

The quantity $m$ in (3.47) is a system constant, which depends on the constructive parameters: $n_1$, $n_2$, $\mu$, $l$ and $A$.

Consider an AC current flowing through the primary, having the amplitude $I_1$ and the angular frequency $\omega$:

$$i_1 = I_1 \sin\omega t \tag{3.48}$$

The open-circuit voltage in the secondary results:

$$e_2 = M_{12} \frac{di_1}{dt} = j\omega M_{12} I_1 \cdot \cos\omega t = j\omega I_1 (m \cdot \cos\alpha) \cdot \cos\omega t \tag{3.49}$$

It can be observed that the amplitude of the induced secondary voltage $e_2$ is proportional to the trigonometric function ($\cos\alpha$):

$$e_2 = E_2 \cos\omega t = K(\cos\alpha) \cdot \cos\omega t \tag{3.50}$$

where $K$ is a proportionality constant.

This dependence evidences the possibility to determine a rotation angle $\alpha$ by measuring the amplitude $E_2$ of the induced voltage.

#### 3.3.7.2 Microsyn

The microsyn is an inductive sensor, which is at the same time associated with electromechanical devices. The microsyn is basically a variable reluctance system that consists of (see Figure 3.25):

- A ferromagnetic rotor, without any windings and rigidly coupled with a target making left/right angular deviations;
- A ferromagnetic stator carrying four orthogonally placed pole pieces, each with two windings.

One coil from every double winding represents the primary coil. These primary inductors are connected in series and are supplied with AC input voltage.

The four secondary coils are connected in such a manner that at the null rotor position the voltages induced in two vis-à-vis looking coils balance the voltages induced in the other opposite coils. The output voltage is equal to zero.

A rotation of the rotor in the clockwise direction increases the reluctance of two opposite coils while decreasing the reluctance in the others, resulting in a net output voltage.

The counterclockwise rotation reverses this effect with a 180° phase shift. A DC direction-sensitive output voltage can be obtained by using phase-sensitive demodulators, for example.

The rotor without any windings and slip rings is a significant constructive advantage. The technical literature underlines the system capability to detect small changes in angles (0.01°) with good sensitivity (5V per degree of rotation). The nonlinearity may vary from 0.5% to 1.0% full scale.

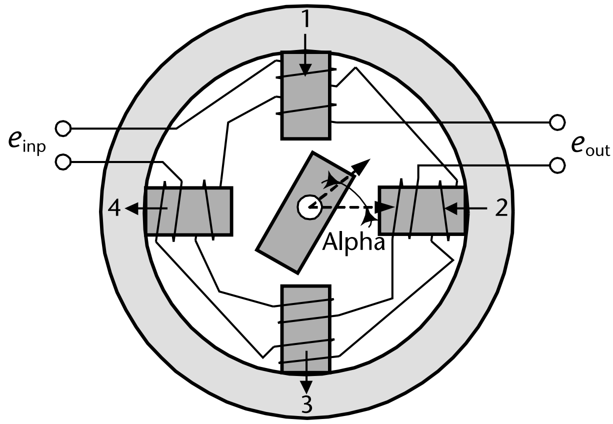


**Figure 3.25** Schematic representation of the microsyn structure.

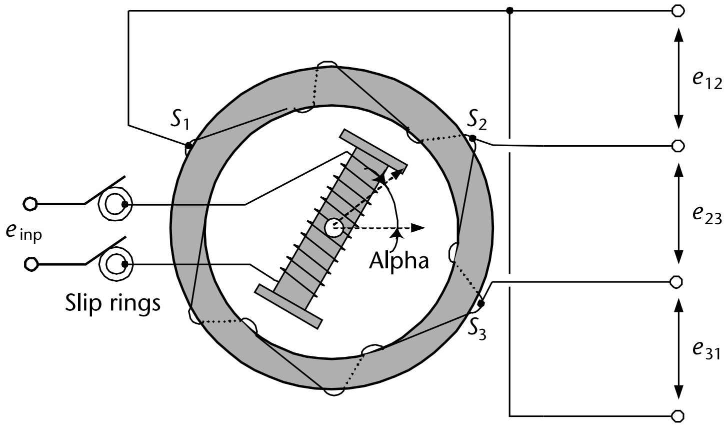


**Figure 3.26** Schematic representation of structure of a synchro and the tree phase system output voltages.

### 3.3.7.3 Synchros

Synchros are also variable reluctance systems. In contrast to microsyn, they have a wound rotor similar to the induction motor and a ring-shaped stator with a set of three phase output coils (see Figure 3.26). These are mechanically spaced by 120°.

The rotor coil is excited by a single phase and its rotation changes the mutual inductance between the rotor coil and the stator coils. The stator coils voltages $e_{12}$, $e_{23}$, and $e_{31}$ having a discrete set of amplitudes for each angular position, can define the rotor position. By interpreting these amplitudes, a table can be established to decode the exact rotor position.

The synchros are primarily used in angle sensing and are commonly applied in control engineering.

### 3.3.7.4 Resolver

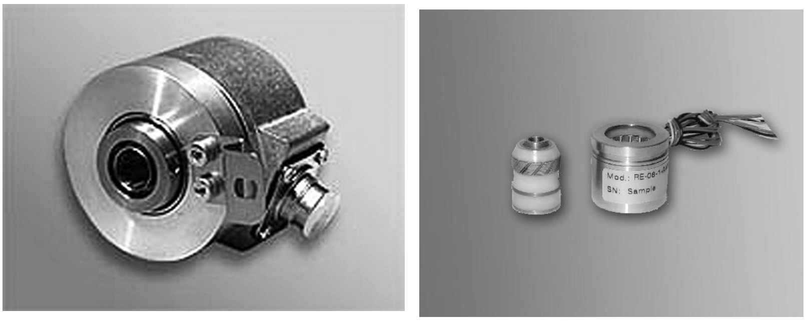
In most applications, resolvers have replaced synchros. Working with a sine and cosine is simpler and requires less conversion and decoding than using three 120° spaced signals. In addition, intelligent integrated electronic circuits for resolver implementation are commercially available (e.g. PGA411-Q1 from Texas Instruments).

This section comments on the main resolver versions used as angle position IS (the rotor effects at most one rotation).

Resolvers are similar in design to the synchros (Figure 3.27). The cylindrical ferromagnetic stator carries two stator windings orthogonally disposed. The ferromagnetic rotor has one or two windings (orthogonally disposed too). Rotor and stator windings are reversible. A pair of windings is excited with AC voltage, the other windings provide induced voltages with angle dependent amplitudes and phases.

Resolvers can be classified into:

- *Sin/cos resolver*;
- *Standard resolver.*



(a) (b)

**Figure 3.27** Several resolver systems. (LTN Servotechnik GmbH).

The sin/cos resolver, also known as resolver with vector decomposition, has only one rotor winding (see its equivalent electrical circuit in Figure 3.28).

The rotor acts as primary winding and the stator contains two secondary windings. For an AC input (see (3.48)), the induced output voltages have the same phase $\omega t$ (see (3.50)):

$$\begin{aligned} e_{S13} &= K(\sin\alpha)\cdot\cos\omega t \\ e_{S24} &= K(\cos\alpha)\cdot\cos\omega t \end{aligned} \tag{3.51}$$

The computation of quotient of both voltage amplitudes allows the computation of trigonometric function ($\tan\alpha$) and hence the determination of the angle *alpha*. Common-mode disturbances such as temperature and deterioration are strongly suppressed or eliminated due to the formation of quotient.

The standard resolver, also known as electrical resolver, with its equivalent circuit in Figure 3.29, has not only two stator windings but also two rotor windings with the same orthogonal arrangement.

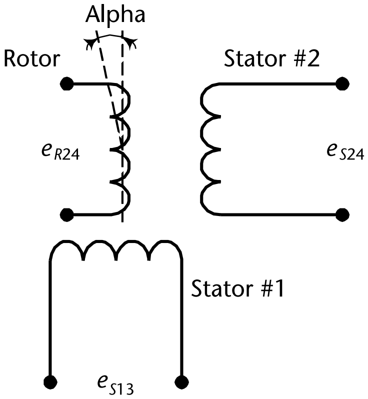


**Figure 3.28** Schematic winding diagram of a sin/cos resolver.

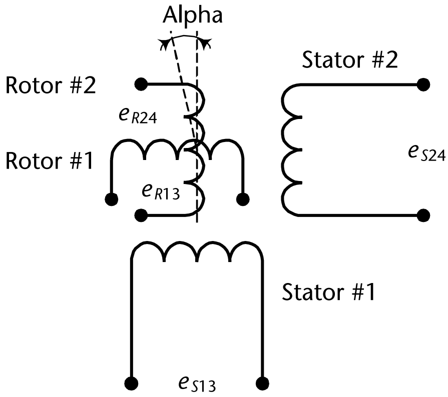


**Figure 3.29** Schematic winding diagram of a standard resolver.

There are several operating modes of standard resolvers in use (see Table 3.2).

The operating mode # 1A designates the first rotor winding to be the system input. This winding is supplied with AC voltage:

$$e_{R13} = E_1 \cos \omega t \tag{3.52}$$

The second rotor winding is short-circuited. The stator windings perform the resolver output and provide two induced output voltages:

$$\begin{aligned} e_{S13} &= \omega K E_1 (\cos \alpha) \cdot \cos \omega t \\ e_{S24} &= \omega K E_1 (\sin \alpha) \cdot \cos \omega t \end{aligned} \tag{3.53}$$

with $\alpha$ (alpha) as the rotation angle between rotor and stator, and $K$ as the proportionality constant. The quotient evaluation is similar to the sin/cos resolver.

**Table 3.2** Summary of Standard Resolver Operating Modes

| *Operating Mode* | *1st. Stator Winding* | *2nd. Stator Winding* | *1st. Rotor Winding* | *2nd. Rotor Winding* | *Functionality* |
|---|---|---|---|---|---|
| # 1A | Both outputs | | Input | Short-circuited | Equivalent to the classical VT |
| # 1B | Input | Short-circuited | Both outputs | | Equivalent to the classical VT |
| # 2 | Both inputs | | Output | | The stator windings are excited with two AC voltages, in phase but with different amplitudes. |
| # 3 | Both inputs | | Output | | The stator windings are excited with two AC voltages, with the same amplitude but having a phase shift of 90°. |

The *operating mode # 1B* is the reversed case. The stator windings are now the system inputs. One of these is supplied with the voltage:

$$e_{S13} = E_1 \cos \omega t \tag{3.54}$$

and the other is short-circuited. In the rotor windings the following output voltages are induced:

$$\begin{aligned} e_{R13} &= \omega K E_1 (\cos \alpha) \cdot \cos \omega t \\ e_{R24} &= \omega K E_1 (\sin \alpha) \cdot \cos \omega t \end{aligned} \tag{3.55}$$

The signal evaluation follows an identical procedure to mode # 1A.

In the operating mode # 2, the inputs (stator windings) are supplied with two synchronous AC voltages (same frequency, same phase) but with different amplitudes, $E_1$ and $E_2$, respectively:

$$\begin{aligned} e_{S13} &= E_1 \cos \omega t \\ e_{S24} &= E_2 \cos \omega t \end{aligned} \tag{3.56}$$

An induced output voltage results in the rotor winding in phase with the inputs:

$$e_{R13} = E_{R13} \cos \omega t \tag{3.57}$$

and whose amplitude $E_{R13}$ has a trigonometric dependence on the difference between the rotation angle $\alpha$ (alpha) and an adjustable offset value $\vartheta$:

$$E_{R13} = \omega K E \cdot \cos(\alpha - \vartheta) \tag{3.58}$$

The electrical advance angle $\vartheta$ is adjustable by means of amplitudes $E_1$ and $E_2$:

$$\vartheta = \operatorname{arctg} \frac{E_2}{E_1} \tag{3.59}$$

The amplitude $E$ in (3.58) has an expression:

$$E = \frac{E_1}{\cos \vartheta} = \frac{E_2}{\sin \vartheta} \tag{3.60}$$

and allows the computation of the rotor angle.

In the operating mode # 3, the inputs (stator windings) are supplied with two AC voltages with the same amplitude and frequency but having a phase shift of 90°:

$$\begin{aligned} e_{S13} &= E \cos \omega t \\ e_{S24} &= E \sin \omega t \end{aligned} \tag{3.61}$$

The output rotor voltage:

$$e_{R13} = \omega KE \cdot \sin(\omega t - \alpha) \tag{3.62}$$

has its phase trigonometrically dependent on the rotation angle $\alpha$.

In this way, the output voltage results time-delayed:

$$e_{R13} = \omega KE \cdot \sin\omega(t - \tau) \tag{3.63}$$

and the time delay $\tau$ is directly proportional to the rotation angle:

$$\tau = \frac{\alpha}{\omega} \tag{3.64}$$

In the advantageous operating mode #3, the measuring of the angle is linear and can be traced back to a phase measuring.

The phase measuring is a counting procedure and normally implies a conversion of the time difference between the zero-crossings of the voltages $e_{S24}$ and $e_{R13}$ into rectangular signal pulse, which forms a gate signal. The counting of high frequency pulses with known and constant frequency starts with zero-crossing of $e_{S24}$ and ends with zero-crossing of $e_{R13}$. The pulses number is a measure of the desired $\alpha$ rotation angle.

A smart synchro behavior can be realized cascading two resolvers.

The first one is a sin/cos resolver and is supplied with AC current. Its rotation angle $\alpha$ represents the target value of the control system for an AC servomotor and is manually adjusted. This resolver generates two output determined voltages with different amplitudes dependent on $\alpha$ (see (3.51)).

These two in-phase voltages supply the second resolver, working in mode #2. The resulting rotor voltage (see (3.58)) is amplified and supplies the AC servomotor. The servomotor shaft drives the resolver rotor (feedback). When the servomotor achieves the desired $\alpha$ value, a rotor voltage equal to zero results, and the motor stops in this position.

A plurality of companies fabricates and commercializes industrial resolvers. For example, LTN Servotechnik GmbH produces capsuled or fully potted compact resolver components (see Figure 3.27).

Resolvers are cost-efficient and extremely rugged. They have a high reliability under harsh environmental conditions such as mechanical shocks and vibrations, are immune to extreme temperature variations and to high toxic stresses or strains caused by chemicals and coolants (compare with Figure 2.7). Up-to-date versions have a considerably longer life expectancy since they do not have electrical slip contacts anymore. The electrical link to the rotor winding is contactless by means of a rotary transformer. In contrast to incremental devices, the resolver delivers the real output, even after an interruption of the supply energy (Section 1.2.3).

Commercial resolvers are medium sized but also miniaturized devices having, shaft diameters: 3 mm to 165 mm and external diameters: 20 mm to 200 mm.

The typical sensitivity is 0.3 V/degree and the error lies below ±5/60 degree (specified for an excitation up to 15V, at a working frequency up to n·10 kHz, and over a large temperature range: –55°C to +155°C).

Beside the classical application, namely rotation angle measurements, resolvers are used for coordinate system transformation (polar in Cartesian and reversed), phase shifting of quadrature signals, and so forth.

### 3.3.8 Final Considerations of Main Inductive Sensor Categories

Following the initial purpose of a gradual introduction and description of the inductive sensors, Section 3.3 ends this succession, which was started with the introductory Chapter 1 and followed up by the specific embedding of the IS into the large world of sensors (see Section 3.2).

The author analyzed and systemized different views in the professional publications and finally proposed to cluster ISs in seven significant types (Section 3.3.1 through 3.3.7). These IS exponents are finally described in Table 3.3. Exceptional cases with low relevance (inductive potentiometer, induktosyn, etc.) have been disregarded.

The magnetoelastic devices are ISs, which exploit the material interaction. The following six types are based on geometry interaction.

Obviously, tasks of ISs could be similarly performed by various types of sensors (capacitive, optical, ultrasonic, etc.). However, there are several fields of application where the use of ISs is imperious, elsewise the single possibility. The sensing of stationary or moving targets, in a range 1 to 100 mm, without a dead zone, with a low energy consumption, independently of material, color, roughness, and so forth remains a very specific job of ISs.

Naturally, they have some weak spots (higher susceptibility to strong external electromagnetic fields, to high temperatures, to metallic environments). Specific measures to improve the IS behavior are described in the next chapters.

**Table 3.3** Summary of the Seven Representative IS Categories

| | | *Self-Generating* | | *Operating Principle* | | | | *Contactless* | |
|---|---|---|---|---|---|---|---|---|---|
| *Book Section* | *System:* | *Active* | *Passive* | *Var. Reluctance* | *Var. Inductance* | *Var. Transformer* | *Eddy Currents* | *-Sensing* | *-Actuation* |
| 3.3.1 | Magneto-elastic | | √ | | √ | | | √ | |
| 3.3.2 | Electrodynamic | √ | | √ | | | | √ | |
| 3.3.3 | Electromagnetic-CML | | √ | | √ | | | √ | |
| 3.3.4 | Electromagnetic-OML | | √ | | √ | | | √ | |
| 3.3.5 | LVDT & RVDT | √ | | | | √ | | √ | |
| 3.3.6 | Eddy currents | | √ | | | | √ | | √ |
| 3.3.7 | Variable transformer | √ | | | | √ | | √ | |

## 3.4 Global Inductive Sensor Market: Size, Share, Growth, Trends, and Forecast

The market of industrial sensors, in general, and of ISs, in particular, is surveyed and analyzed by numerous firms. Reliable sources, which are continuously monitoring and evaluating this market, provide global market research studies that essentially consist of following segments:

- Proximity and Displacement (P&D) sensor type outlook;
- P&D sensor application outlook;
- P&D sensor industry outlook;
- P&D sensor regional outlook;
- P&D sensor key players.

The author makes a summary of such ten extensive reports [11–20] and presents now the sensor market in a top-down hierarchy on three levels:

1. Global sensor market;
2. Global P&D sensor market;
3. Global inductive P&D sensor market.

### 3.4.1 Global Sensor Market

A realistic breakdown of the total worldwide sensor market in 2015 and 2020 by application is illustrated in Figure 3.30 [19].

As already mentioned in Chapter 1, the statistic in Figure 3.30 confirms that the position sensor, with its numerous types and realizations, belongs to the top ten sensors. It will maintain this leading position.

According to [19], its market volume in 2015 was about 7.5 billion U.S. dollars (market share of 8%). At the same time, the forecast predicts a relative stagnation of this market segment; this fact is not valid for the ISs.

### 3.4.2 Global P&D Sensor Market

According to [11], this market is expected to reach 5.32 billion USD by 2020 at a compound annual growth rate (CAGR) of 8.7% from 2015 to 2020 (Figure 3.31).

The automotive and manufacturing industry hold the higher market share (MS) and the Asia-Pacific (APAC) region holds the substantial MS.

A main conclusion refers to the inductive technology. The IS type was, and is, the technological leader. It shows the highest market share of more than 30%.

Comparable values are forecast by [13], namely 6 billion USD by 2022 and a CAGR around 9.5%. The P&D sensor market was valued over 3 billion USD in 2014.

The automotive industry is forecast to reach over 1.9 billion USD in 2022, growing at CAGR of around 10% during 2015 to 2022, and will continue to maintain its leading position. Parking sensor systems is a second large revenue-generating

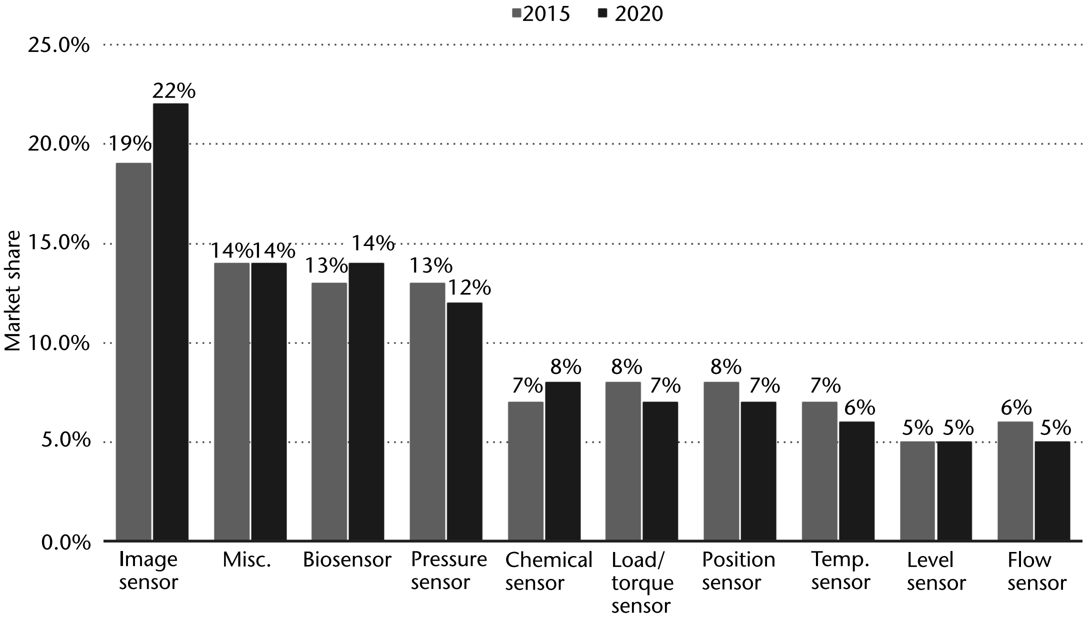


**Figure 3.30** Breakdown of the global sensor market in 2015 & 2020, by application. (Source: Statista / Germany [19], based on BCC research and various sources.)

segment, which is expected to reach over 1.9 billion USD in 2022 with a CAGR of 8% during the same forecast period.

The P&D market is very price sensitive; therefore, regions with low production costs have priority.

The analysis of [20] is more conservative. The global P&D sensor market was valued at 2.93 billion USD in 2013. Europe accounts for more than 30% of the global market revenue share (see Figure 3.32).

The forecast in [20] distantly estimates a CAGR of 2.8% from 2014 to 2020.

The automotive industry has the highest market revenue share contributing with more than 20% to the global market followed by food and beverages. The

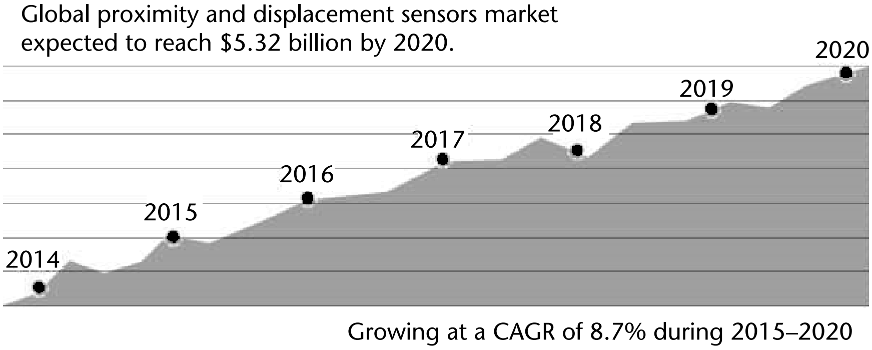


**Figure 3.31** Global P&D sensor market according to [11].

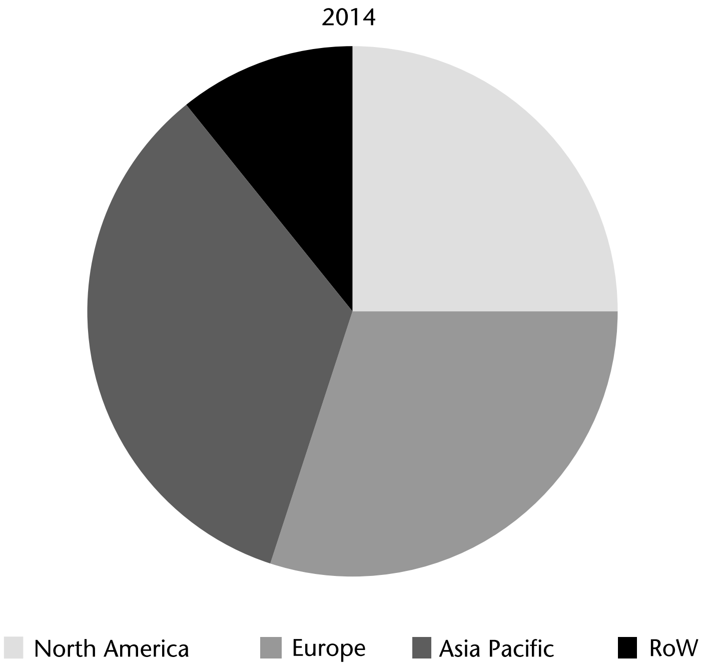


**Figure 3.32** Global P&D sensor market share, by geography, in 2014, according to [20]. RoW = rest of the world.

growth of P&D sensor market is primarily driven by the increasing demand for automated functionalities in manufacturing industries such as food and beverages, pharmaceutical, and process industries among others.

In addition, the lowering cost of components required to manufacture the sensor have further contributed to the growth of P&D.

To complete the information about the global P&D sensor market, Table 3.4 shows a reference analysis including four of the above mentioned research studies.

### 3.4.3 Global IS Market

The study [18] confirms the author's previous affirmations regarding the inductive sensors. As before, IPSs are most widely used in the market, especially for automotive and industrial applications. The sensors are resistant to chemicals and can withstand high temperature and pressure conditions.

In [18] the IPS market is estimated at 1.10 billion USD by 2018, which represents above 20% from the entire market of P&D sensors.

The CAGR is estimated at around 7.09% over the period 2014 to 2020. Increasing demand for automation, travel, and transports are the key drivers, which are making the IPS market to grow lucratively.

Table 3.5 benchmarks four relevant market research reports. Rigorous information offers [16], which provides three different reports for ISs, position sensors, and PSs, respectively. The records in column 3 of Table 3.5 belong to the higher-level report, that is, for ISs.

**Table 3.4** Global P&D Sensor Market

| *Market Segments Based On* | *P&D Market Segments Reported by* | | | |
|---|---|---|---|---|
| | *Allied Market Research [11]* | *GVR Grand View Research [13]* | *Market and Markets [17]* | *Transparency Market Research [20]* |
| P&D Sensor Type Ranking List | Inductive<br>Photoelectric<br>Capacitive<br>Ultrasonic<br>Magnetic<br>LVDT | Inductive<br>Photoelectric<br>Capacitive<br>Ultrasonic<br>Magnetic<br>LVDT | Inductive<br>Capacitive<br>Magnetic<br>Photoelectric<br>Ultrasonic<br>Others | Inductive<br>Photoelectric<br>Capacitive<br>Magnetic<br>Ultrasonic<br>LVDT |
| Application | Parking sensor systems<br>Ground warning system<br>Antiaircraft warfare<br>Assembly line automation<br>Vibration monitoring<br>Roller coasters<br>Conveyor<br>Mobile devices | Parking<br>Ground proximity warning<br>Vibration monitoring and measurement<br>Antiaircraft warfare<br>Roller coasters<br>Conveyer system<br>Mobile devices<br>Assembly line testing | Parking<br>Vibration monitoring and measurement<br>Antiaircraft warfare<br>Roller coasters<br>Conveyer<br>Mobile devices<br>Assembly line testing | Monitoring the distance between objects<br>Metal detection<br>Automobile safety functions |
| Industry | Automotive<br>Pharmaceutical<br>Food and beverage<br>Pulp and paper<br>Elevators and escalators<br>Manufacturing<br>Metals and mining | Automotive<br>Pharmaceuticals<br>Food and beverage<br>Pulp and paper<br>Elevators and escalators<br>Manufacturing<br>Metals mining industry | Aerospace and defense<br>Automotive<br>Food and beverage<br>Pharmaceuticals<br>Consumer electronics<br>Building<br>Industrial manufacturing | Food and beverages<br>Automotive<br>Pharmaceutical<br>Process industries |
| Geography | North America<br>Europe<br>Asia-Pacific (APAC) | North America<br>Europe<br>APAC | North America<br>Europe<br>APAC<br>Rest of the world | North America<br>Europe<br>APAC<br>Rest of the world |
| Key Companies | ifm electronic GmbH<br>Kaman Corporation<br>Keyence Corporation<br>Lion Precision<br>Micron Optics, Inc.<br>Omron Corporation<br>Panasonic Corporation<br>Pepperl + Fuchs GmbH<br>Standex-Meder Electronics Inc.<br>Turck Inc. | Eaton Corporation<br>Omron Corporation<br>Honeywell International Inc.<br>Panasonic Corporation<br>Pepperl+Fuchs GmbH<br>Standex Electronics GmbH<br>Sharp Corporation | Avago Tech. Inc. (Singapore),<br>ifm electronic GmbH<br>Schneider Electric (France)<br>Panasonic Corporation<br>Balluff GmbH<br>Hans Turck GmbH & CO<br>Pepperl + Fuchs GmbH<br>Rockwell Automation<br>Omron Corporation<br>Honeywell International Inc.<br>Fargo Controls<br>Sick AG | Omron Corporation<br>Sharp Corporation<br>Panasonic Corporation<br>among others |

**Table 3.5** Global IS Market and Segmentation of IS Market

| | *Inductive Sensors Market Segments Reported by* | | | |
|---|---|---|---|---|
| *Market Segments Based on* | *HTF-Market Intelligence [14]* | *Market Reports Center [15]* | *Market.biz [16]* | *Mordor Intelligence [18]* |
| IS Type | Variable inductance<br>Variable reluctance sensors<br>AC-operated LVDT sensors<br>Digital I/O LVDT sensors | Cylinder sensors<br>Rectangular sensors<br>Ring and slot sensors<br>Tubular sensors | Self-inductive type<br>Mutual inductive type<br>Eddy current type | Cylindrical ISs<br>Rectangular ISs<br>Slot ISs<br>Ring ISs |
| Industry | Modern machine tools<br>Robotics<br>Avionics<br>Computerized manufacturing | Industrial environments<br>Harsh environments<br>Food industry | Aerospace and defense<br>Automotive<br>Food and beverage<br>Pharmaceuticals | Manufacturing<br>Travel and transport<br>Automotive<br>Defense<br>Energy |
| Geography | China<br>Europe<br>Japan<br>Korea<br>Taiwan | North America<br>Europe<br>China<br>Japan<br>Korea<br>Taiwan | United States<br>China<br>Europe<br>Japan<br>Korea<br>Taiwan | North America<br>Europe<br>APAC<br>Latin America<br>Middle East and Africa |
| Major players | RDP Electrosense<br>ifm efector<br>Micro-Epsilon<br>Trans-Tek<br>Copper Instruments<br>Keyene<br>Comptrol Incorporated<br>Brunswick Instrument<br>Omega Engineering<br>AMETEK Solartron Metrology<br>P3 America<br>Macro Sensors<br>Measurement Specialties<br>American Sensor Technologies | Balluff Inc.<br>Rockwell Automation, Inc.<br>Sick AG<br>Eaton<br>Automation Direct<br>Fargo Controls Inc.<br>Keyence Corporation<br>Pepperl+Fuchs<br>AECO SRL | Omron<br>Pepperl+Fuchs<br>Avago Technologies<br>Schneider Electric<br>Panasonic Corporation<br>Balluff GmbH<br>ifm electronic<br>Rockwell Automation<br>Honeywell International<br>Sick AG<br>Broadcom<br>Eaton<br>Turck Inc.<br>Baumer<br>Carlo Gavazzi<br>Warner Electric (Altra)<br>Proxitron<br>Fargo Controls | 10 leading suppliers":<br>GE<br>Eaton<br>Rockwell Automation<br>Omron<br>Panasonic<br>Freescale<br>Balluff<br>Honeywell<br>Festo<br>Fargo Controls |

## References

[1] Webster, J. G. (ed.), *The Measurement, Instrumentation, and Sensors Handbook*, Boca Raton, FL: CRC Press/IEEE Press, 1999.

[2] Herold, H., *Sensorwirkprinzipien und Sensorsysteme*, Sensortechnik, Heidelberg, Germany: Huethig-Verlag, 1993.

[3] Loos, H. R., *Systemtechnik induktiver Weg- und Kraftaufnehmer*, Ehningen, Germany: Expert Verlag, 1992.

[4] Reuber, C. (ed.), "Band 8: Sensoren und Wandlerbauelemente," *Handbuch der Informationstechnik und Elektronik*, Heidelberg, Germany: Huethig-Verlag, 1989.

[5] http://www.vacuumschmelze.de

[6] Juckenack, D. (ed.), *Handbuch der Sensortechnik–Messen mechanischer Größen*, Landsberg am Lech, Germany: Verlag Moderne Industrie, 1989.

[7] Asch, G., *Les Capteurs en Instrumentation Industrielle*, 7th. ed., Paris, France: Dunod, 2010.

[8] *Maxwell 2014 Training Manual*, Release 15.0, Ansys Inc., Canonsburg, PA, January 2014.

[9] Wilson, J., et. al., *Test and Measurement—Know it All*, Oxford, UK: Elsevier, 2009.

[10] Hauptmann, P., *Sensors—Principles and Applications*, Hertfordshire, UK: Prentice Hall, 1993.

[11] Allied Market Research (www.alliedmarketresearch.com).

[12] ARC Advisory Group (www.arcweb.com).

[13] GVR-Grand View Research (www.grandviewresearch.com).

[14] HTF-Market Intelligence (www.htfmarketreport.com).

[15] Market Reports Center (marketreportscenter.com).

[16] Market.biz (www.market.biz).

[17] MarketsandMarkets (www.marketsandmarkets).

[18] Mordor Intelligence (mordorintelligence.com).

[19] Statista GmbH (www.statista).

[20] Transparency Market Research (transparencymarketresearch.com).

CHAPTER 4

# Inductive Sensing Elements: Evaluation Methods

The descriptions of different ISs in the previous chapters contain references regarding to the kind of primary electrical quantity provided by the sensing element SE of the IS (see Figure 1.2). This intermediate sensor signal results from the primary transduction of sensor input variable (measurand) performed by the SE.

In addition, the graphical representation in Figure 3.1 summarizes the primary electrical quantities resulted from the conversion of nonelectrical measurands detected by sensors, in general, or by ISs, in particular.

This chapter presents analytical, experimental, and numerical techniques to evaluate inductive sensing elements (ISE).

Three procedures are used to provide such evaluations and finally to determine the dependence of its primary electrical quantity on the measurand, for example, travel:

1. *Analytical methods* are the classic way. They sometimes need high mathematical knowledge and provide moderate accuracy. However, they allow one to get a view of the SE operating mechanism.
2. *Measuring* is the traditional option, which offers high precision. However, it requires a physical sample (DUT) of the ISE and corresponding measuring equipment.
3. *Computer simulations* represent the modern up-to-date alternative, which provides all the advantages of the first two methods: good accuracy, no need for advanced mathematics or for hardware, and a scrutiny of the ISE functionality.

These three opportunities are presented in greater detail in the following Sections 4.1, 4.2, and 4.3, respectively.

## 4.1 Analytical Methods of ISE

### 4.1.1 Inductors: Inductance, Impedance, Admittance and Immittance

A list of primary electrical quantities provided by an ISE has to begin with the inductance, which can be named the common denominator of IS.

Inductance is used for the storage of magnetic energy. The storage happens in an electrical device, called inductor (coil), as long as the current *I* keeps flowing through it.

The energy *W* stored in the inductor can be determined by the equation:

$$W = L \cdot \frac{I^2}{2} \tag{4.1}$$

Considering an ideal inductive lumped component, its inductance is defined in Section 3.2. Formulas (3.2) and (3.3) are electromagnetic expressions and allow the computation of the inductance of an air-core coil ($\mu_r = 1$) or a magnetic-core coil ($\mu_r >> 1$) if field quantities are known.

From an electrotechnical point of view, the voltage drop on the inductor and current through it are related via the inductance *L*:

$$v(t) = L \cdot \frac{d}{dt} i(t) \tag{4.2}$$

As a result of the derivation in (4.2), for an ideal, perfect inductor, the current of a sine wave lags voltage by 90°. The dependence between the amplitudes of these harmonic values:

$$V_m = j\omega L \cdot I_m \tag{4.3}$$

defines the impedance $\underline{Z}$ of the perfect inductor:

$$\underline{Z} = j\omega L = jX_L \tag{4.4}$$

where $X_L$ is the inductive reactance and $\omega$ is the angular frequency:

$$\omega = 2\pi f \tag{4.5}$$

The impedance $\underline{Z}$ is essentially a frequency-dependent complex value, which uniquely describes an electrical dipole-network relative to the voltage and current at its terminals. The inductive reactance is purely imaginary. The operator *j* denotes that the inductive reactance dissipates no energy; however, it does oppose current flow.

In some applications, inductors with corresponding inductances $L_i$ are connected in series. The total inductance $L_T$ of *n* such components in series will be increased and, if the inductors are fully noncoupled, is found in the equation:

$$L_T = \sum_{i=1}^{n} L_i \tag{4.6}$$

Logically, it is possible to reduce an inductance connecting inductors in parallel. The total inductance $L_T$ will be less than the value of the lowest inductor:

$$L_T = \frac{1}{\frac{1}{\sum_{i=1}^{n} L_i}} \tag{4.7}$$

Supposing now magnetically coupled inductors, a new term, namely the mutual inductance, has to be defined. This property exists between inductors carrying current when their magnetic lines of force link together.

The mutual inductance $M$ of two inductors with fields interacting can be defined by the equation:

$$M = \frac{(L_A - L_O)}{4} \tag{4.8}$$

where $L_A$ is the total inductance of inductors with fields aiding and $L_O$ is the total inductance of inductors with fields opposing.

Consequently, the coupled inductance $L_T$ of two inductors having inductances $L_1$ and $L_2$ and connected in series with field adding results:

$$L_A = L_1 + L_2 + 2M \tag{4.9}$$

and connected in series with field opposing:

$$L_O = L_1 + L_2 - 2M \tag{4.10}$$

When the two windings are inductively coupled to give transformer action, the coupling factor $k$ is determined by (compare (3.36) in Chapter 3):

$$k = \frac{M}{\sqrt{L_1 \cdot L_2}} \tag{4.11}$$

Real inductors differ from the perfect inductor having a miscellaneous behavior and an equivalent circuit additionally containing a loss resistance $R_S$ and a parasitic parallel capacitance $C_P$ (Figure 4.1). $R_S$ represents the dissipative losses and $C_P$ corresponds to the distributed capacitance between winding turns and inductor terminals.

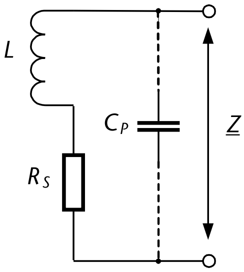


**Figure 4.1** Equivalent circuit of a real inductor.

Supposing the parasitic capacitance can be neglected, the remaining series circuit in Figure 4.1 carries to the general formula of inductor impedance $\underline{Z}$:

$$\underline{Z} = R_S + j\omega L = R_S + jX_L \tag{4.12}$$

Less popular is the reciprocal of the impedance, named admittance $\underline{Y}$, with its real and imaginary part, called conductance $G_S$ and susceptance $B_L$, respectively (see Figure 4.2 with general $\underline{Z}$ and $\underline{Y}$ expressions):

$$\underline{Y} \overset{\text{def}}{=} \frac{1}{\underline{Z}} = G_S + jB_L \tag{4.13}$$

The reason for working with the admittance instead of impedance is to facilitate the computation of parallel LCR networks.

The real parts $R_S$ and $G_S$ represent the losses within the inductor. The imaginary parts $X_L$ and $B_L$ are a measure of the reactive energy stored in the inductor during one period. All these quantities are frequency-dependent, generally.

Beside the Cartesian defining expression of $\underline{Z}$ and $\underline{Y}$ (4.12, 4.13), the corresponding expressions for polar coordinates are very common too:

$$\underline{Z} = |\underline{Z}| e^{j \cdot arctan_Z} \tag{4.14}$$

with magnitude (modulus) $|\underline{Z}|$ and phase angle $arctan_Z$:

$$|\underline{Z}| = \sqrt{R_S^2 + X_L^2} \text{ and } arctan_Z = \arctan\frac{X_L}{R_S} \tag{4.15}$$

The admittance is expressed in polar coordinates by:

$$\underline{Y} = |\underline{Y}| e^{j \cdot arctan_Y} \tag{4.16}$$

with its magnitude $|\underline{Y}|$ and phase angle $arctan_Y$:

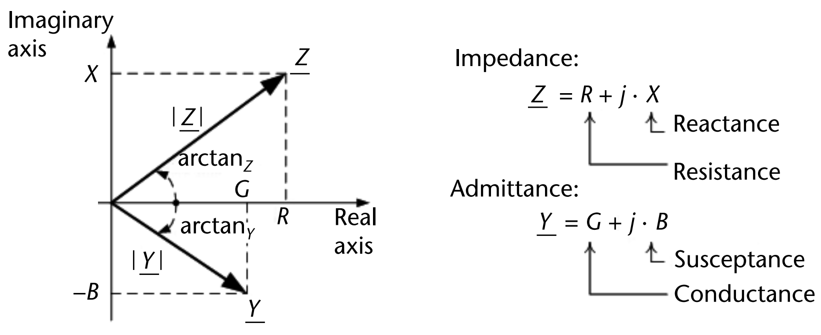


**Figure 4.2** Immittance components dualism: representation of impedance $\underline{Z}$ and admittance $\underline{Y}$ in the complex plane showing the relations between rectangular and polar coordinates. Note that the units are different for each vector.

$$|\underline{Y}| = \sqrt{G_S^2 + B_L^2} \text{ and } arctan_y = \arctan\frac{B_L}{G_S} \tag{4.17}$$

and where:

$$arctan_Y = -arctan_Z \tag{4.18}$$

Based on the display in Figure 4.2, the following relations between rectangular and polar coordinate representation can immediately be deduced:

$$R_S = |\underline{Z}|\cos\left(arctan_Z\right) \text{ and } X_L = |\underline{Z}|\sin\left(arctan_Z\right) \tag{4.19}$$

$$G_S = |\underline{Y}|\cos\left(arctan_Y\right) \text{ and } B_L = |\underline{Y}|\sin\left(arctan_Y\right) \tag{4.20}$$

Since Kirchhoff's laws for voltages and currents also hold for complex quantities, the impedance $\underline{Z}$ of series connection can be expressed by:

$$\underline{Z} = \sum_i \underline{Z}_i \tag{4.21}$$

where $\underline{Z}_i$ is the impedance of the component *i*.

Similarly, the admittance $\underline{Y}$ of a parallel connection results:

$$\underline{Y} = \sum_i \underline{Y}_i \tag{4.22}$$

where $\underline{Y}_i$ is the admittance of the component *i*.

Similar to a series RC-circuit (resistance and capacitance in series), when a DC voltage is applied to a series RL-circuit (resistor-inductor, normally the equivalent circuit of a real inductor) a certain amount of time is required to charge the circuit (compare with Section 1.3.6). The time constant $\tau$ is relevant for the behavior of switched inductive loads (see (10.15)) and is expressed by:

$$\tau = \frac{L}{R_S} \tag{4.23}$$

An important description of the impedance behavior at sweeping frequency is provided by the frequency characteristic of the impedance.

*Example 4.1* To concretely show this characteristic, an inductor with $R_S = 5\Omega$ and $L = 200\ \mu$H is considered. Its impedance is expressed by:

$$\underline{Z} = 5\Omega + j\omega \cdot 2 \cdot 10^{-4}\ \text{H} \tag{4.24}$$

Three traditional graphical representations of the frequency characteristic are normally used:

1. Frequency-response characteristics of the impedance components, in which magnitude and phase angle are apart shown in two distinct graphs (Figure 4.3). Depending on the frequency range extension, the horizontal scale can be linear or logarithmic.
2. Impedance representation in the complex plane (Figure 4.4) is an alternative representation that clearly shows in one chart the variation (rotation and magnitude) of the impedance vector if the frequency *f* varies.

   Ideal resistances *R* are represented by points on the right *x*-axis, ideal inductive reactances $X_L$ on the positive *y*-axis and ideal capacitive reactances $X_C$ on the negative *y*-axis. An impedance that generally consists of resistance and reactance (see (4.12)) is represented by a point on the complex plane. It results an infinite extensive coordinate plane. Figure 4.4 shows the

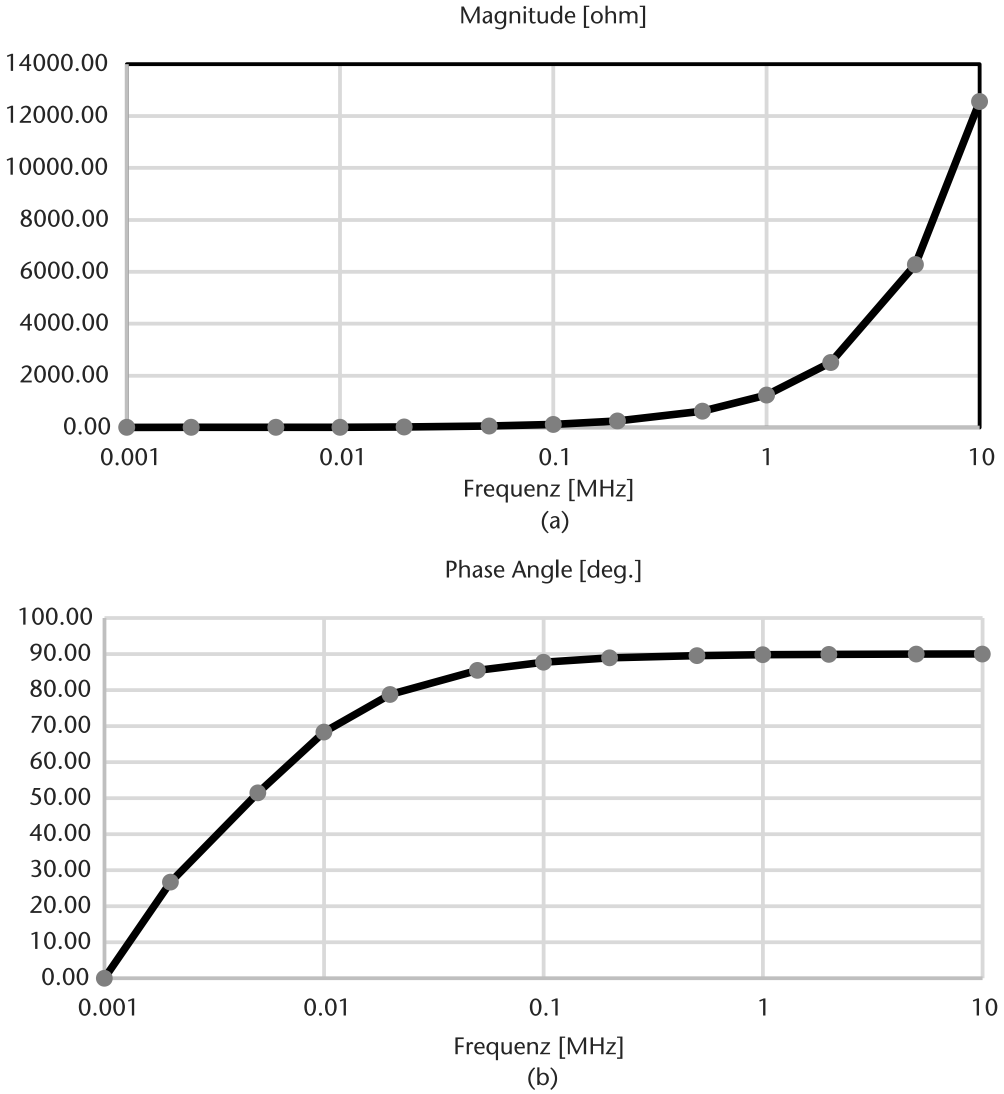


**Figure 4.3** Magnitude (a) and phase angle (b) of the impedance $\underline{Z}$ (4.24) varying with frequency *f*.

frequency dependence of the impedance described by (4.24) for a frequency $f = 1$ kHz to 10 MHz (see the arrow direction). It is a vertical line because in this example the real part $R_S$ is independent of frequency.

The weak point of this method results as the complex plane regions are practically limited. Large real or imaginary values are difficult to be represented.

3. In order o allow the representation of very high resistance and/or reactance values (theoretically unlimited), the orthogonal complex plane is transformed into a circular Smith chart [1]. In terms of mathematical considerations, the chart is a conformal mapping of infinite extensive resistance plane (right Gauss map) into a circle inside (Mobius transformation).

Practically speaking, the following transformation steps lead to the Smith chart (compare with Figures 4.4 and 4.5):

1. The real part $x$-axis (resistance axis) remains unchanged.
2. The $+\infty$ and $-\infty$ extremities of the $y$-axis are moved to link the $+\infty$ extremity of the $x$-axis. The $y$-axis mutates into a main circle having as a diameter the unchanged $x$-axis.
3. Vertical lines corresponding to constant $R$-parameters become circles (reactance circles with $R$ = constant).
4. Horizontal lines corresponding to constant $X$-parameters become arcs of circle (reactance arcs for $X$ = constant).
5. The common link point of all circles and arcs is the right extremity of the main circle diameter ($\infty$ point of all circles and arcs).
6. The origin of the complex plane is zero-impedance $(0 + j0)$ point and the circle center is the reference point $(Z_0 + j0)$.
7. The measuring unit is the unitary (as a rule: ohm).
8. The impedance is represented by a point on Smith chart.

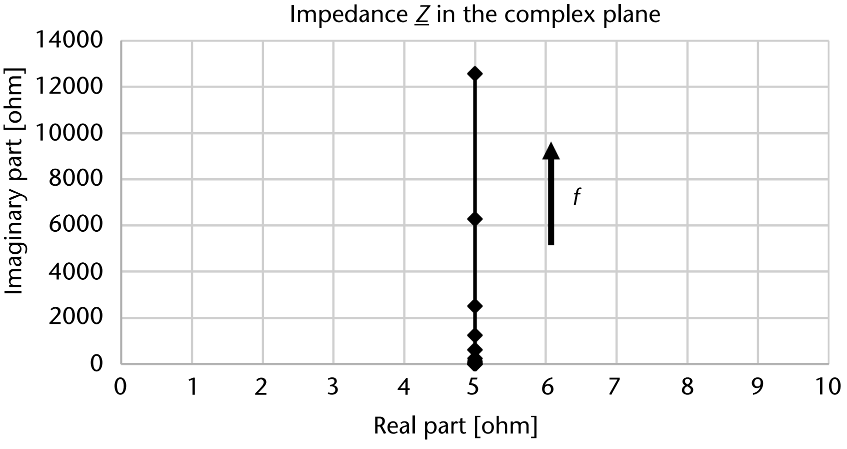


**Figure 4.4** Plots of real and imaginary parts of impedance $\underline{Z}$ (from (4.24)) varying with frequency $f$ represented in the impedance plane.

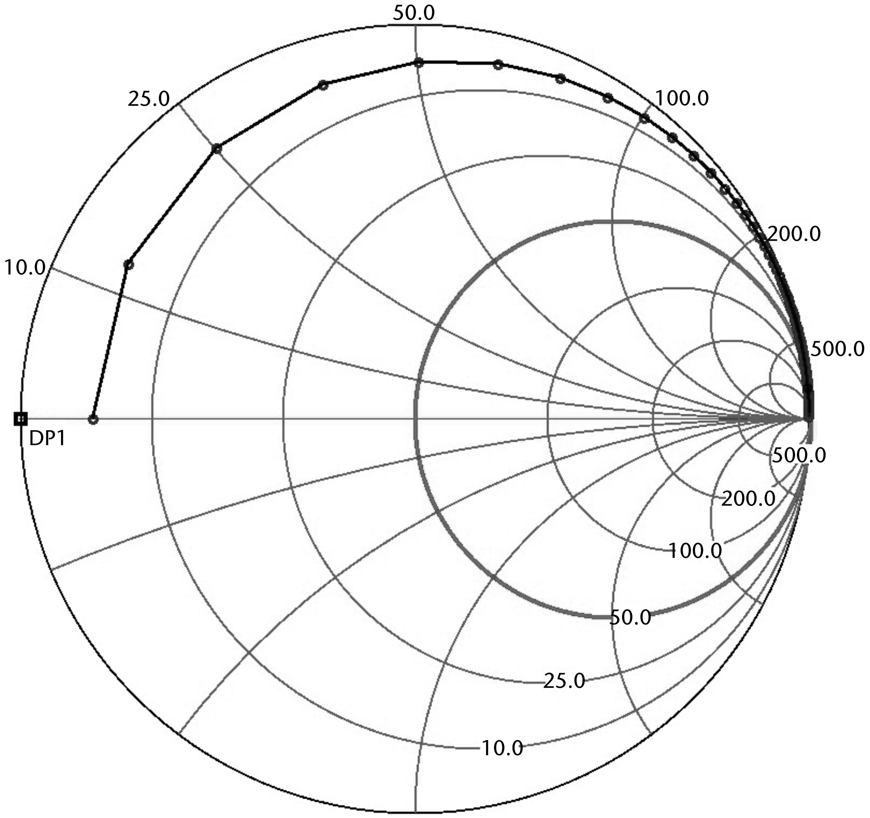


**Figure 4.5** Sweep over frequency of the impedance $\underline{Z}$ (from (4.24)) shown on the Smith chart.

Figure 4.5 shows the frequency dependence of the same exemplifying impedance (4.24) for a frequency range from 0.1 Hz to 10 MHz.

The graph starts very close to the zero-point (resistance is very low: $R_S = 5\Omega$) and runs on the circle diameter (at very low frequency the impedance in quite resistive $\underline{Z} = 5\Omega$). By increasing the frequency, the impedance point evolutes along a trajectory in the upper half of the main circle (range of positive inductive reactances $X_L$) approaching the $+\infty$ point (right extremity of the main circle diameter).

### 4.1.2 Quality Factor of an Inductor

For inductors with losses, the loss angle $\delta$ and the loss factor $D$ are defined by the following formula:

$$D = \tan\delta = \left|\frac{R_S}{X_L}\right| = \left|\frac{G_S}{B_L}\right| \tag{4.25}$$

The inverse quantity is the quality factor $Q_L$ (Q-factor) of the inductor:

$$Q_L \overset{\text{def}}{=} \frac{1}{D} \tag{4.26}$$

For electrical systems, the Q-factor is defined by the ratio between energy stored in reactive system components and the energy dissipated in ohmic resistances, during one time period and multiplied by $2\pi$.

The Q-factor now is the ratio of the inductive reactance to the internal resistance of the inductor. Supposing the real inductor with the equivalent circuit in Figure 4.1 and having the impedance expressed by (4.12), its Q-factor results:

$$Q_L \overset{\text{def}}{=} \frac{\Im \underline{Z}}{\Re \underline{Z}} = \frac{X_L}{R_S} = \frac{2\pi \cdot f \cdot L}{R_S} \tag{4.27}$$

and is affected by frequency $f$, by inductance $L$ (winding dimensions, turns number, core—if any) as well as by losses resistance $R_S$ (DC resistance, core losses, type of wire, etc.). For particular conditions like low working frequency, $R_S$ can be approximated to the coil DC resistance $R_{DC}$.

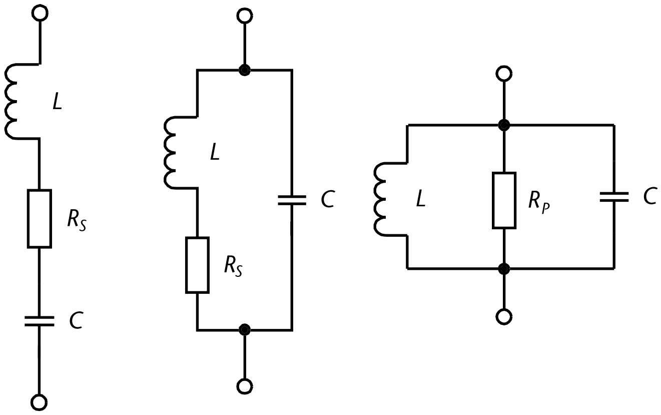
According to (4.27), the Q-factor linearly depends on the frequency. In reality, the $L$ and $R_S$ components of the ISE are measurand dependent but also frequency dependent. Thus (4.27) becomes more complex. Traditional methods to measure the Q-factor are presented in Section 4.2.

### 4.1.3 Impedance and Q-Factor of a Resonant Circuit

The inductor, which represents the main component into the ISE, is frequently not only used as stand-alone but also integrated in a resonant circuit.

Finally, if the inductance evaluation occurs at resonance, their evaluation reduces to an easy and reliable measuring of a real resistance (at resonance, the impedance becomes a real quantity).

To discern which resonant circuit type is more suitable for a certain ISE use, it is advisable to benchmark the classical types: series resonant circuit (Figure 4.6(a))



(a) (b) (c)

**Figure 4.6** Equivalent circuit diagram of a (a) series resonant circuit, and (b) parallel resonant circuit. After a transposition of schematic (b) in pure parallel topology (c), the series loss resistor ($R_S$) appears as a parallel loss resistor ($R_P$).

and parallel resonant circuit (Figure 4.6(b, c)), which contain the inductor with its inductance $L$ and the loss resistance $R_S$ (4.12), as well as a capacitor $C$.

Theoretically, capacitors are also tainted with losses. Normally, capacitors used in ISE have values in the range: 10 pF to 100 nF and are made of ceramic materials. Ceramic capacitors with negligible losses (e.g., NP0-types with $\tan\delta = 1.5\text{E-}3$) are state-of-the-art technology. For this reason, the capacitor $C$ in Figure 4.6 is represented without a parasitic loss resistance.

The components of the total impedance $\underline{Z}$ created in the series resonant circuit by its components (Figure 4.6(a)) can be determined with the following equations:

$$|\underline{Z}| = \sqrt{R_S^2 + \left(X_L - X_C\right)^2} \quad \text{and} \tag{4.28}$$

$$arctan_Z = \arctan\frac{X_L - X_C}{R_S} \tag{4.29}$$

where $R_S$ is the resistance and $X_L$ and $X_C$ are the inductive and the capacitive reactances, respectively. The value $arctan_Z$ is the phase angle, by which current lags voltage in the inductive element or leads voltage in the capacitive element.

With the same notations, the total impedance $\underline{Z}$ created in a parallel resonant LCR circuit with special topology, namely both inductance and series resistance in parallel with capacitance (Figure 4.6(b)) is expressed by the following formulas [2]:

$$|\underline{Z}| = X_C \cdot \sqrt{\frac{R_S^2 + X_L^2}{R_S^2 + \left(X_L - X_C\right)^2}} \quad \text{and} \tag{4.30}$$

$$arctan_Z = \arctan\frac{X_L\left(X_C - X_L\right) - R_S^2}{R_S \cdot X_C} \tag{4.31}$$

When an inductor and a capacitor are connected in series or parallel, they form a resonant circuit. At a specific working frequency called resonance frequency, the inductive reactance equals the capacitive reactance and the resonance effect is installed.

Based on the equality equation:

$$X_L\left(f_0\right) = X_c\left(f_0\right) \tag{4.32}$$

an expression for the resonance frequency $f_0$ can be calculated from. This is the noted Thomson's formula, which gives the resonance frequency with acceptable accuracy:

$$f_0 = \frac{1}{2\pi\sqrt{LC}} \tag{4.33}$$

For resonant circuits having higher losses (inductors with low Q-factor), this formula provides slight inexact results. It can and should be replaced by [3]:

$$f_R = \frac{1}{2\pi}\sqrt{\frac{1}{LC} - \left(\frac{R_S}{L}\right)^2} \tag{4.34}$$

With the same notations like for (4.28), the total impedance $\underline{Z}$ created in the "true" parallel resonant circuit (Figure 4.6(c)) is expressed by the following formulas [2]:

$$|\underline{Z}| = \frac{R_P X_L X_C}{\sqrt{X_L^2 X_C^2 + R_P^2 (X_L - X_C)^2}} \quad \text{and} \tag{4.35}$$

$$arctan_Z = \arctan\frac{R_P(X_L - X_C)}{X_L \cdot X_C} \tag{4.36}$$

Consider the general expression of the impedance of the parallel LCR circuit:

$$\underline{Z} = R + jX \tag{4.37}$$

where the resistance $R$ is purely real and the imaginary part containing the reactance $X$ can have a positive or negative sign. Correspondingly, the parallel LCR network has inductive behavior at given frequencies below resonance frequency $f_0$ or capacitive behavior above this characteristic frequency (see Section 8.1.2).

When the resonant circuit is used to work at its own resonance frequency $f_0$ (self-excitation), the relation (4.32) is fulfilled and the impedance $\underline{Z}$ is purely ohmic. This value is low at the series equivalent circuit (equal to $R$) and high resistive at the parallel equivalent circuit. The substantial higher value of the parallel topology is the main decision criterion to favor this in electronic implementations.

Moreover, the schematic in Figure 4.6(b), which is a physical representation of an inductor ($L$ and $R_S$ components) connected in parallel to a capacitor ($C$ capacitance), can be lightly transposed into an equivalent pure parallel topology (Figure 4.6(c)) that is easy to handle. The new equivalent parallel resistance $R_P$ is purely ohmic and can be expressed by:

$$R_P = \frac{1}{R_S} \cdot \frac{L}{C} \tag{4.38}$$

Similar to $R_S$, $R_P$ is an easily measurable quantity (see Section 4.2).

For practical application and design activities, Table 4.1 summarizes several expressions of the $L$-$C$-$R_S$-$R_P$-$Q$ relationships. Additionally, Table 8.1 shows the electrical magnitudes (voltage and current) of the resonant LCR circuit.

Finally, when the aforementioned losses of the capacitor $C$ are negligible, the quality factor $Q$ of the resonant LC circuit at the resonance is equal to the quality factor of the inductor with losses $Q_L$.

*Example 4.2* Reconsidering the exemplary inductor with $L = 200\ \mu\text{H}$ and $R_S = 5\Omega$ (see (4.24)) and connecting a capacitor with $C = 1$ nF in parallel to this, a high Q-factor resonant circuit with specific parameters results:

**Table 4.1** Summary of Equations Characterizing Resonant LCR Circuits

| *Quantity* | *Expressions* |
|---|---|
| Resonance frequency | $f_R = \frac{1}{2\pi}\sqrt{\frac{1}{LC} - \left(\frac{R_S}{L}\right)^2} \approx f_0 = \frac{1}{2\pi}\sqrt{\frac{1}{LC}}$ |
| Quality factor of the resonant circuit | $Q = \frac{\omega_R \cdot L}{R_S} = 2\pi f_R \cdot \frac{L}{R_S}$<br>$Q = \frac{2\pi f_R L}{R_S} \quad Q = \frac{R_P}{2\pi f_R L}$<br>$Q = \frac{1}{2\pi f_R C R_S} \quad Q = 2\pi f_R C R_P$<br>$Q = \frac{1}{R_S} \cdot \sqrt{\frac{L}{C}} \quad Q = R_P \cdot \sqrt{\frac{C}{L}}$ |
| Loss resistances | $R_P = \frac{(\omega_R \cdot L)^2}{R_S} = \frac{(2\pi f_R)^2 \cdot L^2}{R_S}$<br>$R_S = \frac{1}{Q} \cdot \sqrt{\frac{L}{C}} = \frac{R_P}{Q^2} \quad R_P = Q \cdot \sqrt{\frac{L}{C}} = R_S \cdot Q^2$ |

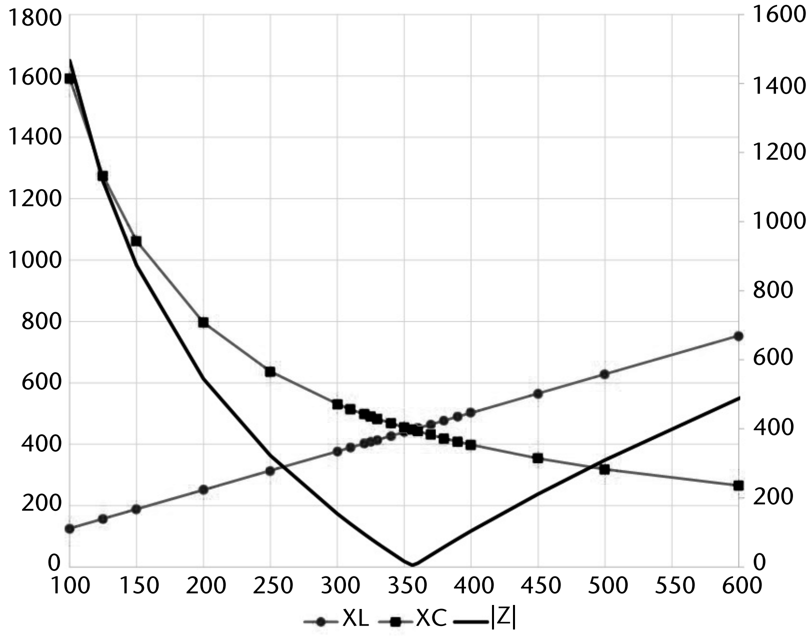
- Resonance frequency $f_R \approx 356$ kHz (see (4.34));
- Quality factor $Q_L = 89.442$ (see Table 4.1);
- Equivalent parallel resistance $R_P = 40$ kΩ (see (4.38)).

Connecting these inductors and capacitors in series results in a series resonant circuit with the same resonance frequency. Figure 4.7 benchmarks these resonant circuits regarding their parameters: impedance magnitude $|\underline{Z}|$ as well as the inductive and capacitive reactances $X_L$ and $X_C$. The impedance dip of the series circuit and impedance peak of the parallel circuit at the resonance frequency $f_R \approx 356$ kHz can immediately be seen.

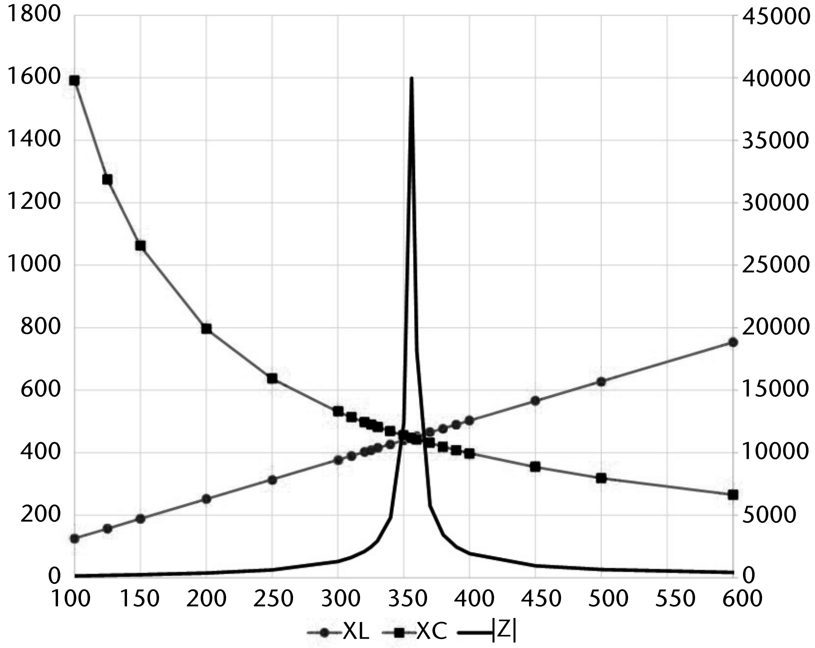
An interesting-looking picture is illustrated in Figure 4.8. On an impedance Smith chart, the frequency sweep circle of the parallel resonant circuit joints at $f_R$ the significant chart point, namely the reference point (resistance circle set for 40 kΩ) and is tangent to the reactance arcs $X_L = 20$ kΩ and $X_C = 20$ kΩ, respectively.

The resonant circuit is very often used in ISE. There are indeed two operation modes:

1. Normally, the circuit is the resonator of an electronic oscillator and determines the working frequency. That means the resonant circuit is used at resonance and the aforementioned considerations are valid.
2. The resonant circuit can be driven by an external oscillator at an excitation frequency, which is different from its resonance frequency.



(a)



(b)

**Figure 4.7** Frequency dependence of the impedance magnitude $|\underline{Z}|$ (right axis) as well as of the inductive and capacitive reactances $X_L$ and $X_C$ for: (a) series resonant circuit (see Figure 4.6(a)), and (b) parallel resonant circuit (see Figure 4.6(c)). The unit for both vertical axes is ohm.

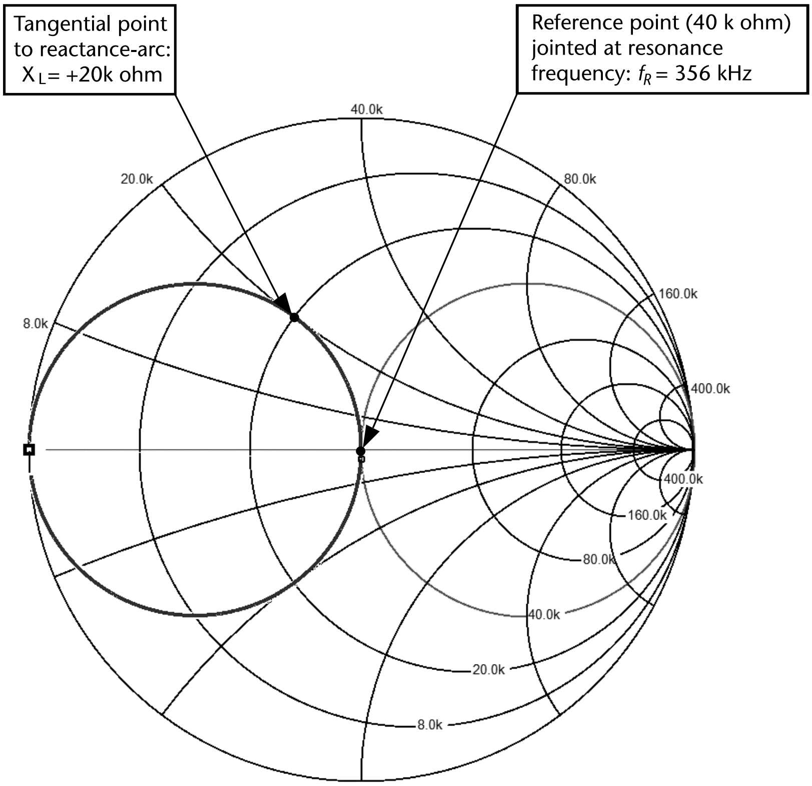


**Figure 4.8** Smith chart: Frequency sweep diagram of the impedance $\underline{Z}$ of the parallel resonant circuit (see Figure 4.6(c)).

The measure of how the resonant circuit responds to driving signals of a frequency lateral to the resonance frequency is the bandwidth $B_Q$. It is a frequency range having the limits defined by a 3 dB reduction of the impedance relative to the resonance. At these frequency limits, the reactive part of the impedance is equal to the real part. The impedance drops to 70.7 percent peak-impedance $(1{:}\sqrt{2})$. Solving this mathematical condition for the lateral frequencies $f_1$ and $f_2$ two expressions result, which show their dependence on the resonance frequency $f_0$ (4.33) and Q-factor.

Finally, the bandwidth $B_Q$ results:

$$B_Q \overset{\text{def}}{=} f_2 - f_1 = \frac{f_0}{Q} \tag{4.39}$$

It is advisable to use the resonant circuit inside its bandwidth.

## 4.2 Measuring Methods to Evaluate the ISE

The measuring of the sensing inductor's parameters ($\underline{Z}$, $L$, $R_S$ and/or $R_P$) is an important task during the ISE design activity as well as for the finalization of the

IS electronics, which has to evaluate the primary quantity provided by ISE. Hence, the measuring methods can be classified in:

- Metrological methods;
- Methods suitable to be implemented in IS.

### 4.2.1 Experimental Methods

When designing an ISE it is necessary to know its characteristics, firstly the Q-factor and the equivalent parallel resistance $R_P$ in dependence on the sensing distance, over the entire sensing range.

The application report [4] simply and clearly describes three modern procedures:

1. A vector network analyzer (VNA) can measure the complex impedance of the inductor under test (see (4.12)) over a suitable range of frequency and show this dependence in a Smith-chart. Placing the cursor at the targeted resonance frequency, a frequency selection occurs and the parameters $L$ and $R_S$ are displayed. The searched values Q and $R_P$ can be calculated by means of Table 4.1. The network analyzer uses an automatic error correction to eliminate the effect of couplers and junctions. Because of that, a preliminary calibration procedure with standard, very precise terminations is necessary. The method provides very accurate results but it requires the most expensive investment, unfortunately.
2. Similarly, an impedance analyzer could directly provide the value $R_P$.
3. The report [4] finally describes a traditional method that uses a signal generator and an oscilloscope. The practicable method was many times exercised by the author and the results were amazingly accurate. The procedure requires a capacitor in addition to the inductors (see Figure 4.9), drives the resonant circuit at its resonance frequency $f_R$, and makes use of the pure real value of the impedance $Z(f_R) = R_P$. Therefore, the $R_P$ is measured at $f_R$ only. It is an iterative method consisting of the following steps:
    1. A sine-wave signal is set at the generator output (preferred low impedance output) and kept constant;
    2. The signal frequency is adjusted until the resonance occurs ($v_{LC}$ displayed on the scope screen reaches maximum);
    3. The variable resistor $R$ (preferred low-inductance type) is adjusted so that $v_{LC} = 1/2 \cdot v_{GEN}$;
    4. The steps above are repeated to get better accuracy;
    5. The final R-value measured with a classical ohmmeter represents the $R_P$ value.

The topic of this subsection is the object of a large presentation in [5]. It classifies the methods in three basic groups:

1. Current and voltage methods based on impedance determination. They are suitable for all type of inductors and use vector voltmeters to determine the

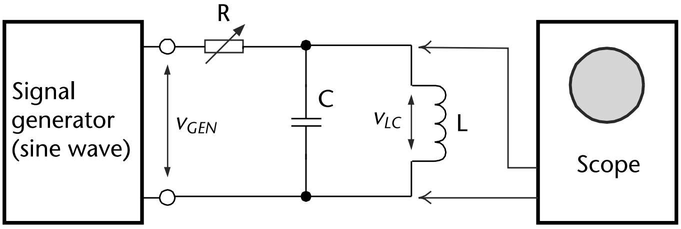


**Figure 4.9** Classical method using a signal generator and oscilloscope to measure the equivalent $R_P$ of a resonant circuit [4].

magnitude and phase angle of the impedance or apply the three-voltmeter method [5]. With a remarkable effort on hardware, the methods can be today considered as obsolete for measurements on the laboratory scale.

2. Bridge and differential methods based on comparison of unknown impedance with a reference impedance until a state of balance is reached. The volume [5] describes four significant bridge types: Maxwell-Wien, Hay, Carey-Foster, and bridge with Wagner-branch. The choice of the most suitable configuration depends on the properties of the unknown element.

   The frequently used Maxwell-Wien and Hay topologies have the unknown inductor in the first leg, fix resistors in two additional legs and a parallel or series RC circuit with adjustable elements in the fourth leg. In addition, [5] mentions the high number of other existing AC bridges and the option of a universal bridge that can be configured for different applications by switching elements.

   Alternating current bridges are cost-efficient laboratory devices, which can measure from DC up to the order of a few hundreds of MHz with very high precision and sensitivity. They need only a signal generator and a zero indicator in the diagonal branch.
3. Resonance methods are based on application of a series or parallel resonant LC circuit (Figure 4.6), where the unknown inductor is coupled with a variable very low-loss capacitor $C_V$ (Figure 4.10).

Using the inductor as part of a resonant circuit and tuning $C_V$ to maximum voltage $v_C$, the Q-factor can be directly measured (see voltage ration in (4.40)). The unknowns $R_S$ and $L$ are obtained from this formula and from the resonance condition applied for the test frequency:

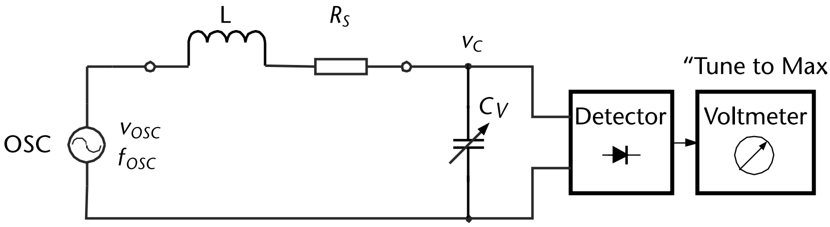


**Figure 4.10** Unknown inductor ($L$, $R_S$)as part of a resonant circuit to determine its own quality factor and inductance (Q-meter principle). Variable capacitor $C_V$ is a test device.

$$Q = \left|\frac{V_{\mathrm{Cmax}}}{V_{\mathrm{OSC}}}\right| = \frac{1}{2\pi f_{OSC} \cdot R_S C_V} \text{ and } L = \frac{1}{\left(2\pi f_{OSC}\right)^2 \cdot C_V} \tag{4.40}$$

where $f_{OSC}$ is already known as the frequency of the supplying oscillator.

The unknown inductor could be the element of a measuring bridge circuit driven again at resonance. One leg of the Wheatstone bridge contains the inductor to be measured ($L$ and $R_S$) in series with the variable capacitor $C_V$, the other three legs have noninductive resistors $R_2$, $R_3$ and $R_4$ (Figure 4.11).

The adjustment consists of two steps: provide resonance state by varying the capacitor $C_V$ and balance the bridge ($v_{\mathrm{DB}} = 0$) using the resistors. Under these conditions, the parameters of the unknown inductors results in:

$$L = \frac{1}{\left(2\pi f_{OSC}\right)^2 \cdot C_V} \text{ and } R_S = \frac{R_2 \cdot R_4}{R_3} \tag{4.41}$$

where $f_{OSC}$ is the already known frequency of the supplying oscillator.

Beside this easiest bridge with resonant circuit, [5] describes more complex topologies having more legs with resonant circuits.

In conclusion, these experimental methods perform a measuring if the bridge is already balanced and/or if the resonant circuit is in resonance. That requires manual or automatic adjustments of tuning and reference elements to align the output to zero (bridge balancing) or to maximal value. The impedance to be determined results from comparing it to final reference values.

Obviously, these procedures need measuring hardware and instruments and—above all—require adjustment activities, which cannot be easily implemented in a sensor.

### 4.2.2 Measuring Methods Suitable to Be implemented in IS

Bridge circuits and resonant circuits remain base elements to be implemented in ISs in order to make sensing functionality possible. Compared to previous experimental applications, their mode of use and the generation method of the output are now different.

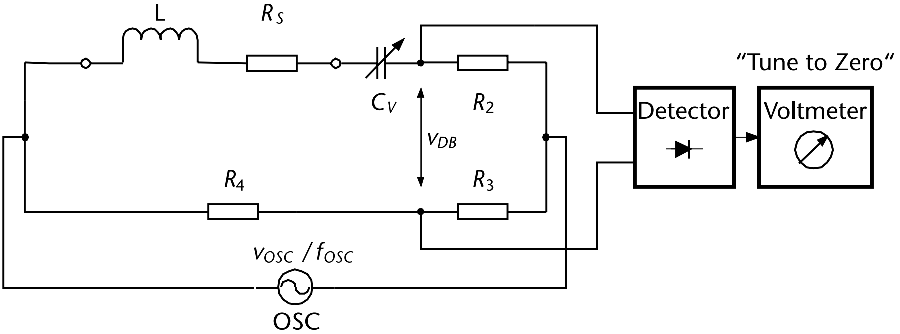


**Figure 4.11** Circuit diagram applied in resonance method.

From now on, these electrical circuits work in balanced state only at zero-measurand position. Apart from this particular case, they will be unbalanced and the unbalance magnitude provides indirect information about the measurand.

Some implementations are already described in Chapter 3. Nevertheless, the present section gives an entire overview of the popular methods:

*Bridge operations*. They are used in combination with differential sensing elements, consisting of two inductors to capture the measurand (e.g. Figures 3.10, 3.16). They form a bridge-branch. The second branch consists of two resistances. At the sensor trimming after fab-out, the resistors are adjusted to provide a bridge output voltage equal to zero for the rest position of the armature or plumber.

*Resonant circuits applications*. The target movement acts on the steady state of the LC network, which corresponds to the zero or rest position of the target, and changes the parameters of the voltage $v_{LC}$ over resonant circuit; that is, amplitude, frequency, or phase deviation.

Resonant circuit applications are preferentially applied in IPS. The inductor that forms the ISE of these sensors is connected in parallel to a fix-value capacitor. The inductor parameters $L$ and $R_S$ (see Figure 4.12) are influenced by the target and have to be evaluated in the sensor. The resulting resonant LC circuit is supplied by

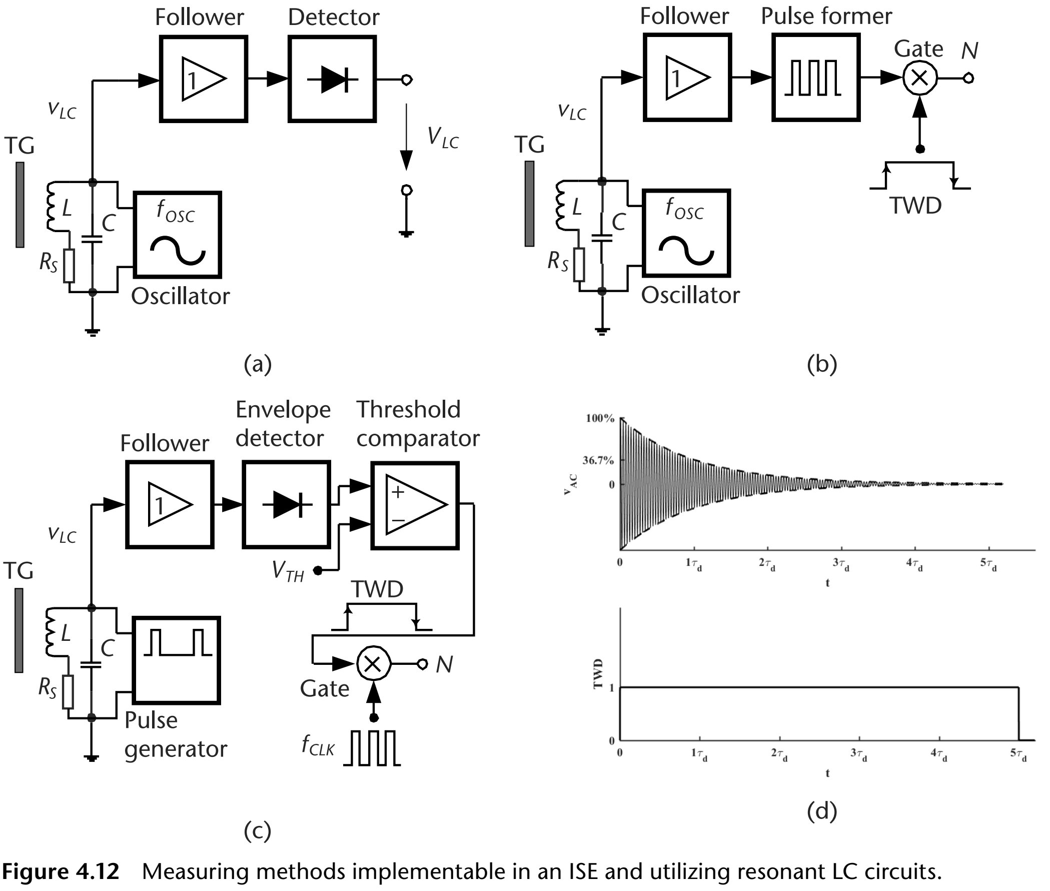

Follower
Detector
$v_{LC}$
TG
L
C
$R_S$
$f_{OSC}$
Oscillator
$V_{LC}$
(a)
Follower
Pulse former
Gate
N
$v_{LC}$
TG
L
C
$R_S$
$f_{OSC}$
Oscillator
TWD
(b)
Envelope detector
Threshold comparator
Follower
$v_{LC}$
TG
L
C
$R_S$
+
–
$V_{TH}$
TWD
Gate
N
Pulse generator
$f_{CLK}$
(c)
100%
36.7%
0
$v_{AC}$
0
$1\tau_d$
$2\tau_d$
$3\tau_d$
$4\tau_d$
$5\tau_d$
t
TWD
1
0
0
$1\tau_d$
$2\tau_d$
$3\tau_d$
$4\tau_d$
$5\tau_d$
t
(d)

**Figure 4.12** Measuring methods implementable in an ISE and utilizing resonant LC circuits.

a free-running oscillator [7], whose working frequency is determined by the resonance frequency $f_R$ of this LC circuit (see (4.34)).

The classic evaluation method is an AC to DC signal conversion (see Figure 4.12(a)). The evaluation channel picks up the oscillating signal $v_{LC}$ by means of a follower impedance convertor (high impedance connection):

$$v_{LC}(t) = V_{lcm}(|\underline{Z}|) \cdot \sin(2\pi f_R \cdot t) \tag{4.42}$$

and performs a signal amplitude demodulation (peak detector, average value detection, PLL detection, etc.)(see Section 9.2). The resulted DC voltage $V_{LC}$ has an ideal proportionality to the amplitude $V_{lcm}$, which in turn depends on the inductor impedance $|\underline{Z}|$. The output voltage $V_{LC}$ is an electronic image of the contactless actuation of the target.

At the same time, the variation of parameters $L$ and $R_S$ causes a drift of the resonance frequency $f_R$ (see (4.34)). This influence can also be used to evaluate the ISE. The sinusoidal oscillation $v_{LC}(t)$ with the frequency $f_R$ is converted in rectangular pulse waves having the same frequency (see Figure 4.12(b)). The measuring of frequency occurs in a counter circuitry. The base schematic contains a gate unit, which permits the rectangular pulses to go through as long as the time window signal TWD is HIGH. This enabling signal with an accurate period is provided by a control unit and controls the second gate input. The number $N$ of the run-through pulses gives the resonance frequency $f_R$.

The method shown in Figure 4.12(c) is a little bit different and it is essentially an indirect measuring method of the Q-factor. The procedure takes advantage of the behavior of resonant LC circuits supplied by periodically short current pulses. The signal $v_{LC}(t)$ is now a decaying harmonic wave with the frequency $f_R$, whose envelope follows an exponential function with defined decay time $\tau_d = 2L/R_S$:

$$v_{LC}(t) = V_{lcm} \cdot e^{\frac{-R_S}{2L} \cdot t} \cdot \cos(2\pi f_R t) = V_{lcm} \cdot e^{\frac{-\pi f_R}{Q} \cdot t} \cdot \cos(2\pi f_R t) \tag{4.43}$$

This time depends on the Q-factor of the resonant circuit and its measuring gives the electronic image of the target contactless actuation too.

The LC circuit in Figure 4.12(c) is now supplied from a pulse generator (pulse duration as short as possible and cycle frequency much lower than the frequency $f_R$) and the measuring chain begins with an envelope detection followed by a threshold comparator. Its digital output signal represents a time window TWD whose duration is to be measured. The signal goes to the second enable input of the gate unit (reversed situation compared to Figure 4.12(b)). The main gate input is supplied with rectangular clock pulses with accurate cycle period $T = 1/f_{\mathrm{CLK}}$. The number $N$ of the run-through pulses gives the Q-factor. The window can be large enough to wait until the amplitude of $v_{LC}(t)$ reaches the zero value (comparator threshold equal to zero). That happens after five decay times $\tau_d$ (see Figure 4.12(d)). The Q-factor can be calculated with:

$$Q = \frac{\pi}{5} \cdot f_R \cdot T \cdot N \tag{4.44}$$

To speed up the measuring, the counting procedure can be stopped after a decay time $\tau_d$ when the envelope theoretically achieves 36.7% from its initial maximal peak. The threshold of the comparator $V_{TH}$ has to be set for this level. A small number $N_1$ of counted pulses results and the Q-factor is now:

$$Q = \pi \cdot f_R \cdot T \cdot N_1 \tag{4.45}$$

*Example 4.3* Returning to exemplifying inductor with $R_S = 5\Omega$ and $L = 200\ \mu\text{H}$ (see (4.24)) connected in parallel to a capacitor $C = 1$ nF, the resulting LC circuit has a resonance frequency $f_R = 356$ kHz (see (4.34)) and a quality factor $Q_L = 89.442$. Consider, for example, rectangular clock pulses with cycle period $T = 1\ \mu\text{s}$ ($f_{\text{CLK}} =$ 1 MHz). For a specific decay time $\tau_d = 80$ µs, the resulting number of pulses will be $N = (5 \cdot 80\ \mu\text{s})/1\ \mu\text{s} = 400$.

Substituting these values in (4.44), the result is an evaluated value $Q_L = 89{,}427$, which is very close to the real value of 89,442. Obviously, the offered accuracy is lower because of the quantization. Nevertheless, the method is cost-effective, suitable for microcontroller applications and requires lower supply current.

The combination of the above-described versions; namely two resonant circuits integrated in a bridge arrangement, is also possible [8].

Finally, the measuring of an induced voltage remains also a possibility to evaluate an ISE measurand. The voltage can be generated as a result of a provoked induction (see Figure 3.3) or by means of a transformer (see Figure 3.4).

## 4.3 Modern Computer-Assisted Analysis and Synthesis of ISE

The conclusion in Section 3.3.3, after the analytical evaluation of the system shown in Figure 3.9, emphasizes the limits of such common procedures and anticipates the only possibility to overcome their obstacles is using computer-aided engineering (CAE)-based field simulation tools. This conclusion is the best introduction to the section.

In order to understand these modern computer-aided engineering methods, it is instructive to briefly review the principles of the electrodynamics.

### 4.3.1 Fundamentals of Computer-Aided Electromagnetic Field Simulation

The scientific and technical publications quite unanimously use traditional graphical and/or analytical methods to model and describe the operating principle of the various ISs.

De facto, this task is an electromagnetic field solving problem and today there are powerful hardware and software tools that effectively and efficiently can analyze, synthesize, and optimize all IS systems.

Having significant experience in this field and believing in this opportunity, the author's intention is to release a first book predominantly based on computer-assisted analysis instead of using mathematical methods.

The theory behind the field solvers originally belongs to the brilliant physicist James Clerk Maxwell. His key conclusions in electromagnetic fields, that is, a time-varying magnetic field generates a spatially varying electric field and a time-varying electric field produces a spatially varying magnetic field, are analytically expressed by means of fundamental equations of the electromagnetic theory, namely Maxwell's equations, established in 1873. In point form, they are [2]:

$$\begin{aligned} \nabla \times \mathbf{E} &= \frac{-\mathrm{d}\mathbf{B}}{\mathrm{d}t} \\ \nabla \times \mathbf{H} &= \mathbf{J} + \frac{\mathrm{d}\mathbf{D}}{\mathrm{d}t} \\ \nabla \cdot \mathbf{D} &= \rho_v \\ \nabla \cdot \mathbf{B} &= 0 \end{aligned} \tag{4.46}$$

The electromagnetic quantities in these formulas (bold characters for vectors and normal for scalars) and their measuring units are specified in the nomenclature section at the end of this book. The quantities are functions of position and time (position vector **p** and time *t*).

In (4.46), the symbol $\nabla$ represents the vector operator (called del operator or nabla symbol) and acts on vectors via cross product (the curl) or via dot product (the divergence), respectively [2].

Maxwell's equations are supported by the following auxiliary equations:

Constitutive relations, which have a simple form for an isotropic medium:

$$\begin{aligned} &\mathbf{D} = \varepsilon \mathbf{E} \quad \text{with } \varepsilon = \text{medium's permittivity} \\ &\mathbf{B} = \mu \mathbf{H} \quad \text{with } \mu = \text{medium's permeability} \end{aligned} \tag{4.47}$$

Electrical current density **J** expressions:

$$\begin{aligned} &\mathbf{J} = \sigma \mathbf{E} \quad \text{with } \sigma = \text{medium's conductivity (Ohm's law)} \\ &\mathbf{J} = \rho_v \mathbf{V} \quad \text{with } \rho_v = \text{volume charge density} \end{aligned} \tag{4.48}$$

Polarization **P** and magnetization **M** of linear materials:

$$\begin{aligned} &\mathbf{P} = \chi_e \varepsilon_0 \mathbf{E} \quad \text{with } \chi_e = \text{electric susceptibility} \\ &\mathbf{M} = \chi_m \mathbf{H} \quad \text{with } \chi_m = \text{magnetic susceptibility} \end{aligned} \tag{4.49}$$

Extended equations for displacement **D** and magnetic flux density **B**:

$$\begin{aligned} &\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P} \quad \text{with } \varepsilon_0 = \text{absolute permittivity of free space} \\ &\mathbf{B} = \mu_0 (\mathbf{H} + \mathbf{M}) \quad \text{with } \mu_0 = \text{absolute permittivity of free space} \end{aligned} \tag{4.50}$$

Even for powerful systems, solving the field problem that means the computing of the above-mentioned equations for a given topology (structure and materials), excitations and boundary conditions could be a difficult, at least a nonconvergent task. For this reason, the commercial simulation programs contain solver software modules, which are specialized in particular field problems. A ranking list according to the relevance for ISE includes the following modules:

- *Eddy current field solver* computes the oscillating magnetic field that exists in a structure due to a distribution of AC currents. It also computes current densities, taking into account all eddy current effects (including skin and proximity effects). An impedance matrix, force, torque and current flow may also be computed from the energy stored in the magnetic field and from energy losses in conductors.
- *Axial eddy current field solver* computes the oscillating magnetic field that exists in a structure by external oscillating magnetic fields. It also computes current densities, taking into account all eddy current effects (including skin effects). Current flow may also be computed from the energy stored in the magnetic field and from energy losses in conductors.
- *Magnetostatic field solver* computes the static magnetic field that exists in a structure given a distribution of DC currents. An inductance matrix, force, torque and flux linkage may also be computed from the energy stored in the magnetic field.
- *Electrostatic field solver* computes the static electric field that exists in a structure given a distribution of DC voltages and static charges. A capacitance matrix, force, torque and flux linkage may also be computed from the energy stored in the electric field.
- *AC conduction field solver* computes the AC currents that flow in a dielectric given a distribution of AC voltages. An admittance matrix and current flow may also be computed from the energy stored in the electric field and from energy losses in dielectrics.
- *DC conduction field solver* computes the DC currents that flow in a dielectric given a distribution of DC voltages. An admittance matrix and current flow may also be computed from the energy stored in the electric field and from energy losses in dielectrics.

The algorithm of the eddy current field solver restricts Maxwell's equations to time-varying currents. Such currents flowing in a conductor produce a time-varying field in planes perpendicular to the conductor. In turn, this magnetic field induces eddy currents in the source conductor and in any other conductor parallel to it. The solver assumes that all time-varying electromagnetic quantities have the form:

$$F(t) = F_m \cos(\omega t + \vartheta) \tag{4.51}$$

where $\omega$ is the angular frequency at which all quantities are oscillating (4.5).

Using the aforementioned equations and Euler's formula, the first two Maxwell's equations reduce to:

$$\begin{aligned} \nabla \times \mathbf{E} &= -j\omega \mathbf{B} \\ \nabla \times \frac{1}{\mu}\mathbf{B} &= \sigma \mathbf{E} + j\omega\varepsilon \mathbf{E} \end{aligned} \tag{4.52}$$

The auxiliary quantity that the field simulator actually solves for is **A**, the magnetic vector potential [6]. It is a help quantity without physical signification and is defined by:

$$\nabla \times \mathbf{A} = \mathbf{B} \tag{4.53}$$

Substituting this into the second part of (4.52), the result is:

$$\nabla \times \frac{1}{\mu}(\nabla \times \mathbf{A}) = \mathbf{E}(\sigma + j\omega\varepsilon) \tag{4.54}$$

A solution of (4.54) for **E** in terms of **A** is given by:

$$\mathbf{E} = -j\omega\mathbf{A} - \nabla\Phi_E \tag{4.55}$$

where $\Phi_E$ is the electric scalar potential. Substituting the right side of (4.55) into (4.54) results in the final field equation used by the field simulator to systematically compute the field quantities beginning with **A**:

$$\nabla \times \frac{1}{\mu}(\nabla \times \mathbf{A}) = (\sigma + j\omega\varepsilon)(-\nabla\Phi_E - j\omega\mathbf{A}) \tag{4.56}$$

The right side of (4.56) is in the form of a complex conductivity multiplied by the complex value of **E**. It is therefore equal to the complex current density **J**. Correspondingly, the total current $I_T$ is the sum of three components:

1. The source current density due to the differences in electric potential:

$$I_s = -\int \sigma\nabla\Phi_E \, dv \tag{4.57}$$

2. The induced eddy current density due to time-varying magnetic fields:

$$I_e = -\int j\omega\sigma\mathbf{A}dv \tag{4.58}$$

3. The displacement current density due to time-varying electric fields:

$$I_d = \int j\omega\varepsilon(-j\omega\mathbf{A} - \nabla\Phi_E)dv \tag{4.59}$$

The eddy and displacement components are a function of frequency and become increasingly significant as the frequency increases.

Induced currents allow magnetic fields to penetrate conducting media only to a certain depth $\delta$ called skin depth, which is approximated by the formula:

$$\delta = \sqrt{\frac{2}{\omega\sigma\mu_0\mu_r}} \tag{4.60}$$

where the quantities are aforementioned.

Currents will be concentrated near the surface of the conductor (skin effect), decaying rapidly pass the skin region (exponential function). The formula above indicates $\delta$ gets smaller as the frequency increases.

In addition to the field simulation, an eddy current field solver can perform the computation of the impedance matrix that summarizes the relationship between AC voltages and AC currents in each loop of multiconductor systems (a practical example is the coil of an inductive sensing element). For a system having $n$ current loops (e.g., coil turns), the relationship would be expressed in matrix form by an $(n \times n)$ impedance matrix:

$$\left[\underline{V}_k\right] = \left[\underline{Z}_{ik}\right]\left[\underline{I}_i\right] \tag{4.61}$$

where:

$\underline{V}_k$ and $\underline{I}_i$ are the voltage and current phasors ($i = 1$ to $n$);

$\underline{Z}_{ii}$ is the self-impedance of the loops 1 to $n$, expressed as a complex function of the AC resistance and reactance:

$$\underline{Z}_{ii} = R_i + j\omega L_{ii} \tag{4.62}$$

$\underline{Z}_{ik}$ is the mutual impedance between loops $i$ and $k$ ($i$, $k = 1$ to $n$) and depends on the corresponding mutual inductances $M_{ik}$:

$$\underline{Z}_{ik} = \underline{Z}_{ki} = j\omega M_{ik} \tag{4.63}$$

The field simulator breaks down the impedance matrix computation into two parts. First, it solves for the inductance matrix associated with the model ($L$-matrix). It then solves for the resistance matrix ($R$-matrix), and finally, it combines both.

The computation of the inductance of a loop bases upon the calculation of the system average energy and leads to the final formula [7, 8]:

$$L = \frac{\mu}{I_m^2} \iiint \mathbf{H} \cdot \mathbf{H}^* \, dv \tag{4.64}$$

where $I_m$ is the amplitude of the loop AC current and $\mathbf{H}^*$ is the complex-conjugated vector of the magnetic field strength $\mathbf{H}$, expressed by means of its real and imaginary parts:

$$\mathbf{H}^* = \Re\{\mathbf{H}\} - j\Im\{\mathbf{H}\} \tag{4.65}$$

The computation of the AC resistance is based on the calculation of the ohmic loss and leads to the final formula [7, 8]:

$$R = \frac{1}{\sigma \cdot I_m^2} \iiint \mathbf{J} \cdot \mathbf{J}^* \, dv \tag{4.66}$$

where $I_m$ is the amplitude of the loop AC current and **J** and **J***  are the electric current densities: vector and complex-conjugated vector, respectively.

### 4.3.2 Field Simulation Software Tools

There are a significant number of world leader companies in computer-aided engineering software for electromagnetic design and analysis. They provide similar but still different integrated software tools for 2D and/or three-dimensional (3D) geometries, which perform magnetostatic, electrostatic, time-harmonic, and other simulations for a wide variety of applications including sensors.

The software tools make the computation, simulation and optimization of electromagnetic field problems possible. The problem solving occurs in three main steps:

1. *Preprocessing*, which basically includes the geometric modeling of the system to be analyzed, definition of the simulation region and its boundaries, the materials- and excitations-setup, fixing of the solution type, and so forth.
2. *Real solving*, representing the main activity of the tool. As already mentioned, the solver is structured in modules having a given solving functionality, that is, electrostatic, magnetostatic, eddy-currents, and so forth.
3. *Postprocessing*, which ends the simulation task, providing the calculated values of the field quantities in analytical and/or graphical display form.

The theoretical base is already described in Section 4.3.1. To solve these equations, the entire simulation region containing different objects and/or free spaces is divided into a very large number of 2D/3D elements (triangles and tetrahedrons) applying the finite element analysis (FEA) or other advanced procedures [9].

This section presents some important provider of such tools in an arbitrary order. The majority were established between 1980 and 1990 and continuously improved their products.

Opera FEA simulation software (Illinois, USA) [10] currently develops six supported analysis packages: TOSCA, ELEKTRA, SCALA, CARMEN, SOPRANO and TEMPO. ELEKTRA analyzes time dependent electromagnetic fields, including the effects of eddy currents, in three dimensions. There are 3 analysis options: the time variation can be transient, for steady state, or analyze with eddy currents, which can be induced in moving conductors with a specified linear or rotational velocity in the presence of a static field.

INTEGRATED engineering software (Manitoba, Canada) [11, 12] provides a large spectrum of solver modules essentially developed for 2D or 3D applications, respectively: ELECTRO or COULOMB for electrostatic fields, MAGNETO or AMPERES for magnetostatic fields, and OERSTED or FARADAY for time-harmonic fields (including eddy-currents). In contrast to Opera tools, these solvers are based on the boundary element method (BEM) [12].

CST Computer Simulation Technology AG (Darmstadt, Germany) [13] provides the CST STUDIO SUITE®, which is an electromagnetic simulation software package with solvers for a wide of applications across the electromagnetic spectrum. They are from static and low frequency to microwave and RF and for a large range of applications. The CST package comprises a comprehensive spectrum of modules: CST MICROWAVE STUDIO®, CST PARTICLE STUDIO®, CST CABLE STUDIO®, CST MPHYSICS STUDIO® and, last but not least, the CST EM STUDIO® (CST EMS). This is an easy-to-use tool for the design and analysis of static and low-frequency EM applications such as motors, sensors, actuators, transformers, and shielding enclosures.

### 4.3.3 Simulation with ANSYS Maxwell Tool

The presentation of electromagnetic field simulation providers continues now with a significant global player, namely the large company ANSYS, Inc. (Pennsylvania, USA). Besides a large spectrum of simulation tools (for a breadth of technologies: fluid mechanics, structural mechanics, etc.), ANSYS naturally offers software for electromagnetic problems [14]. They are called Maxwell (the classical, an initial Ansoft product) and HFSS (for microwave applications).

The simulations in this book are made with the Maxwell software package. The author and his collaborators have been using this tool for more than 20 years with positive experience and close results to the reality.

Maxwell performs FEA to simulate electromagnetic fields [15, 16]. It solves the electromagnetic field problems by solving Maxwell's equations (see (4.46)) in a finite region of space with appropriate boundary conditions and user-specified initial conditions in order to obtain a solution with guaranteed uniqueness. A suitable set of equations is applied based on the solver selected by user; that is, magnetostatic, eddy current, magnetic transient, and electrostatic.

The eddy current solver is the most important for inductive sensors. It solves steady state, sinusoidally varying magnetic fields in frequency domain. Additionally, it solves linear materials and considers displacement currents. The source of magnetic field can be sinusoidal currents in conductors or time-varying external magnetic fields. The induced fields, skin, and current proximity effects are also considered.

Static magnetic fields caused by DC currents and permanent magnets (e.g., IS in Figure 3.3) can be solved by the magnetostatic solver, which can consider linear and nonlinear materials.

The finite element method refers to a numerical technique from which the solution is numerically obtained from an arbitrary geometry by breaking it down into simple pieces called finite elements. Maxwell solution methods are comparatively described in [9].

Overall, there are three Maxwell options: RMxprt, 3D, and 2D. The software supports 3D but also 2D geometric models. The finite elements are corresponding tetrahedrons or triangles. RMxprt is an interactive tool used for designing and analyzing electrical machines and is suitable for variable transformers (see Section 3.3.7).

The interactive dialogue user-machine occurs by means of a versatile graphical user interface (GUI), which contains:

- Three bars, the menu, tool, and status bar;
- Several windows: modeler and project tree windows (for example, see Figure 3.12) as well as project manager, properties, message, and progress windows;

- Coordinate entry window (Cartesian, cylindrical, or spherical).

In order to save computer simulation time and money, it is recommended to use 2D models, if possible. This modeling remarkably reduces the solving effort without any significant influence on the result accuracy.

The majority of inductive sensors have a rotationally symmetric coil. For this reason, a modeling in the *rz* plane is well suitable (see Figure 4.13). The 2D geometric model is a cross section into the coil and the solver considers a full rotation (Theta $\vartheta = 0$ to 360°) around the vertical *z*-axis.

Consequently, the field vectors with the following components are computed:

The first computed vector **A** having a perpendicular position on the RZ plane and a value dependent on the position in this plane:

$$\mathbf{A}(\mathbf{p}) = A_{\vartheta}(r,z) \cdot \mathbf{e}_{\vartheta} \tag{4.67}$$

The derived electric field quantities having similar position:

$$\mathbf{E}(\mathbf{p}) = E_{\vartheta}(r,z) \cdot \mathbf{e}_{\vartheta} \tag{4.68}$$

$$\mathbf{J}(\mathbf{p}) = J_{\vartheta}(r,z) \cdot \mathbf{e}_{\vartheta} \tag{4.69}$$

In contrast, the derived magnetic field quantities are represented by vectors in the RZ plane:

$$\mathbf{H}(\mathbf{p}) = H_R(r,z) \cdot \mathbf{e}_R + H_Z(r,z) \cdot \mathbf{e}_Z \tag{4.70}$$

$$\mathbf{B}(\mathbf{p}) = B_R(r,z) \cdot \mathbf{e}_R + B_Z(r,z) \cdot \mathbf{e}_Z \tag{4.71}$$

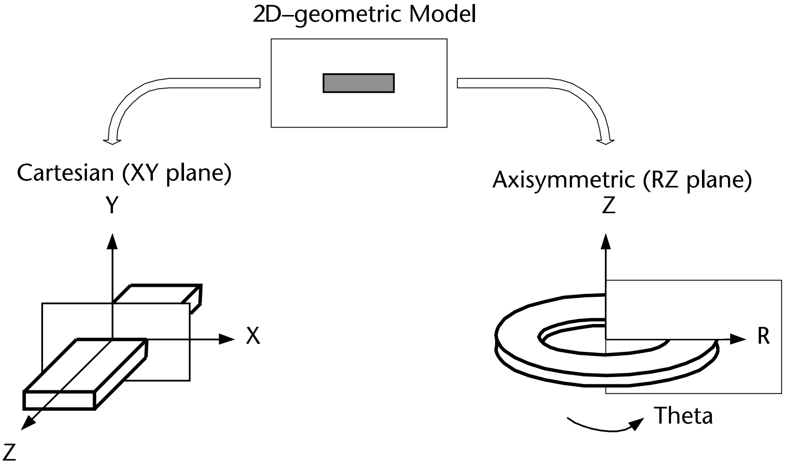


**Figure 4.13** Available Maxwell coordinate systems for 2D solution types and corresponding models.

The vectors $\mathbf{e}_R$, $\mathbf{e}_Z$, and $\mathbf{e}_\vartheta$ are the corresponding unit vectors. They are present in the usual expression for the axis-symmetric representation of general vector **X**, defined by the position vector **p**:

$$\mathbf{X}(\mathbf{p}) = X_R(\mathbf{p}) \cdot \mathbf{e}_R + X_Z(\mathbf{p}) \cdot \mathbf{e}_Z + X_\vartheta(\mathbf{p}) \cdot \mathbf{e}_\vartheta \tag{4.72}$$

Finally, the solver integrates the vector **A** to calculate the currents:

$$I = \int \mathbf{J} \cdot dS \tag{4.73}$$

where $S$ is the area over which the current flow is computed and **J** is the current density vector, given by:

$$\mathbf{J} = (\sigma + j\omega\varepsilon)(-\nabla\Phi_E - j\omega\mathbf{A}) \tag{4.74}$$

For *xy* models (see Figure 4.13), the area $S$ is found by sweeping the current flow line represented in the *xy*-plane into *z*-direction—forming a 3D surface. The computed current flow is the current per meter depth in the *z*-direction. For *rz* models, the area $S$ is found by revolving the flux line represented in the *rz* plane 360° around the *z*-axis, forming a 3D surface. The computed current flow is the total current that passes through this surface.

### 4.3.4 Flow Chart of a Maxwell Field-Simulation Project: Concrete Example

After the formulation of the CAE basics, after the introduction of the Maxwell software and its particular aspects in the previous sections, the present section gives a detailed, systematic presentation of a simulation project. To be in step with actual practice, the items in the Maxwell flowchart will be explained, making parallel references to a real, practical example.

The selection of the appropriate and available resources, such as solver type and solution type, as well as coordinate system, is a higher-level project item. In our example, they are eddy current solver, 2D solution type, and Cartesian coordinates system.

A Maxwell field simulation project starts with preprocessing steps, which are manually done by the user (see Figure 4.14).

The first action is the preparation of the geometrical simulation model with its belonging entities (step #1). The model can be drawn with the Maxwell graphic editor. Maxwell can also import CAD files (AutoCAD, CATIA, etc.). The real material has to be allocated to each object. Maxwell possesses a large default material library, which cannot be modified by the user. Users can generate material personal libraries and change and/or complement the initial material specification (see Figure 4.15).

*Example 4.4* This refers to the ISE of an IPS, model IA-M12 (see Section 2.1.1.6 for model definition and Figure 3.20 for the SE schematic representation).

The *rz* geometrical model in Figure 4.16 of this practical example slightly differs from the schematic representation in Figure 3.20 and consists of five parts from the following materials with the following characteristics (see Figure 4.17):

1. The ISE coil having 100 turns from enamel-insulated copper wire with a diameter of 0.1 mm. Copper material parameters: relative permeability $\mu_r = 0.999991$ and specific conductivity $\sigma = 5.8E7$ S/m.
2. Coil former made from Teflon that is a good insulator and has nonmagnetic properties ($\sigma = 0$ S/m and $\mu_r = 1$).
3. Ferrite pot core, type Sch9/F08 (manufactured by Neosid/Germany) with $\mu_r = 700$ and $\sigma = 1$ S/m. It is important to mention that the core has a central through bore (typical diameter of 2 mm).
4. *Cylindrical housing* from chromed brass with $\mu_r = 0.99999$ and $\sigma = 1.59E7$ S/m;
5. *Protection* and *insulation cap* between core coil and cylindrical housing also made from Teflon.

The definition of the simulation region and boundary conditions (step #2) requires theoretical users' knowledge and practical experience. The simulation region is necessary to be defined in order to specify the room in which FEM calculations are carried out. This region should completely enclose the model and should have sufficient clearance from geometry. It should be sufficiently large because the solver considers no field in the area outside this region. On the other hand, the larger the simulation region, the higher the computing effort and solving time.

Boundary conditions define the behavior of the magnetic field at the border of the simulation region. The common type, namely balloon models, defines the space outside the simulation region as being infinitely large. Magnetic flux lines are neither tangential nor normal to the boundary.

Assign setup conditions (step #3) refer to the specification and operating mode of the testee. The excitation current (value, phase, direction) and its density in conductors (stranded for uniform distribution or solid for eddy distribution) are assigned (in our example, a current of 2 mA peak-to-peak flows through all 100 turns that are solidly wired and correspondingly solid configured).

It is necessary to evaluate the SE losses provoked by eddy currents in metallic system parts, and so, the processing with eddy currents consideration must be unconditionally enabled. The solver considers proximity effects between parts as well as skin effects in every component carrying currents. It calculates specific skin depths in these components and sees to the corresponding current density (see (4.60)).

The define solution setup (step #4) requests fixing the following features (values in brackets are from Example 4.4):

- Maximum number of passes (iterations): default is 10 (reference example: 20 passes).
- Adaptive working frequency (reference example at 500 kHz) and—optionally—frequency range limits and sweep step if a frequency sweep is desired.
- Convergence criteria has a minimum number of passes and minimum converged passes.
- Accuracy defined by the percent energy error—default is 1% (for reference example: 0.1%). For the solver, some fundamental defining equation provides an error evaluation for the solved fields. Energy produced by these error terms is computed and compared with the total energy. The result is called percent

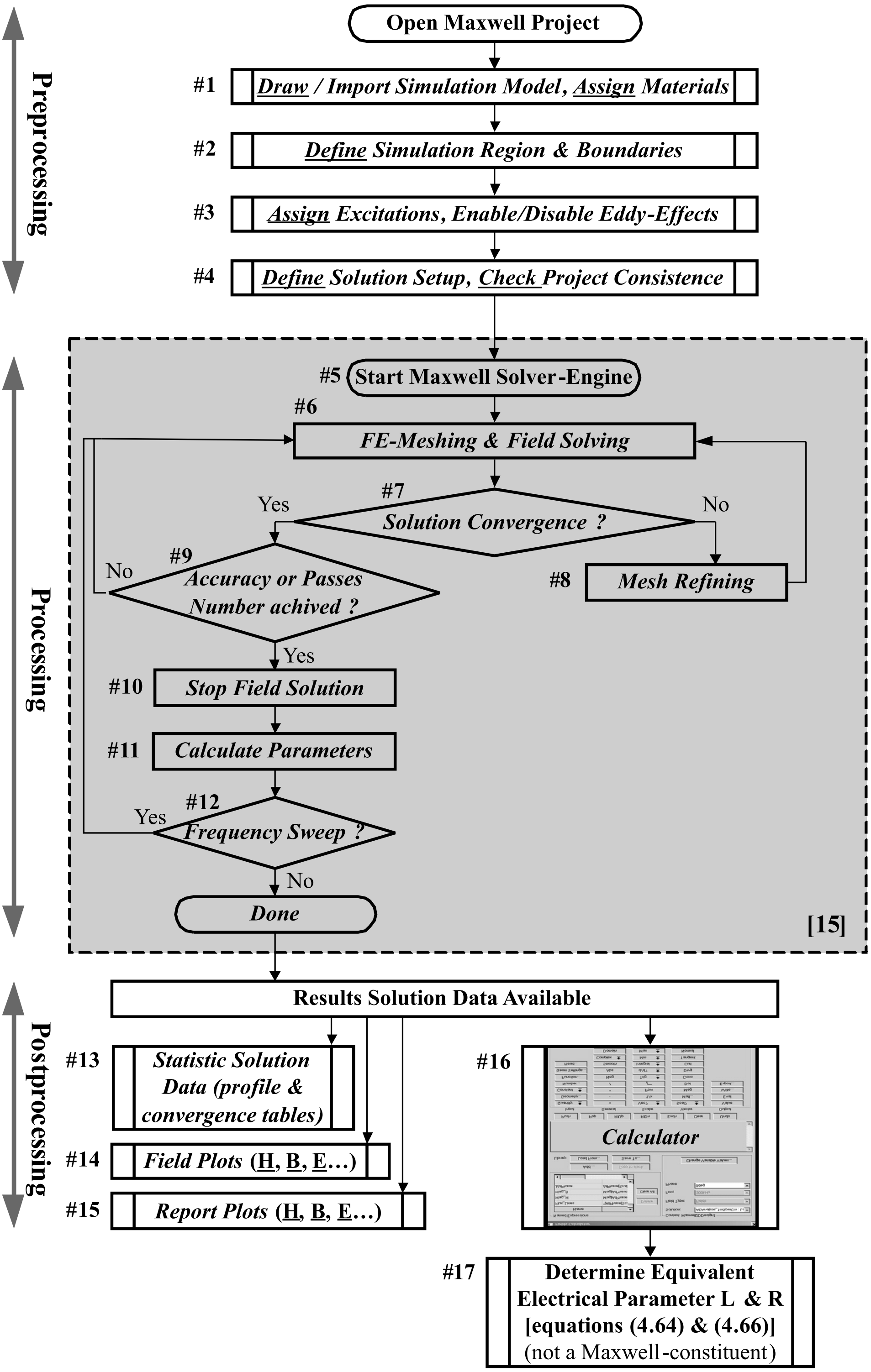
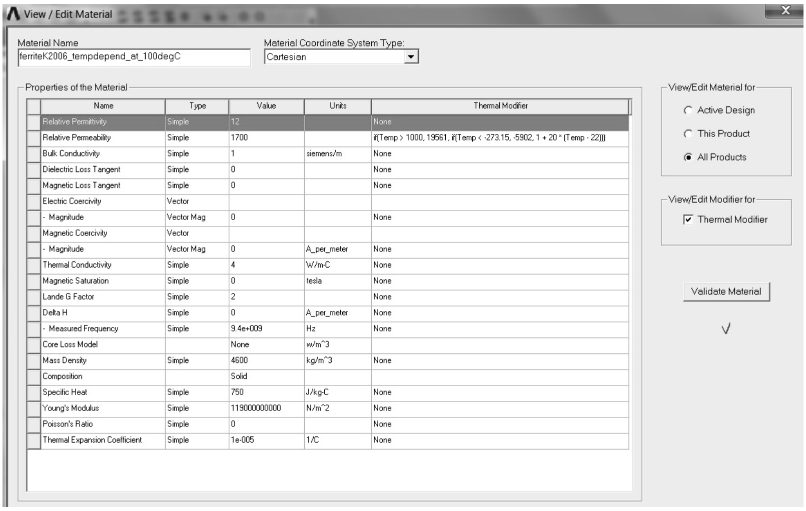

Preprocessing
Processing
Postprocessing
Open Maxwell Project
#1 Draw / Import Simulation Model, Assign Materials
#2 Define Simulation Region & Boundaries
#3 Assign Excitations, Enable/Disable Eddy-Effects
#4 Define Solution Setup, Check Project Consistence
#5 Start Maxwell Solver-Engine
#6 FE-Meshing & Field Solving
#7 Solution Convergence ?
Yes
No
#8 Mesh Refining
#9 Accuracy or Passes Number achived ?
No
Yes
#10 Stop Field Solution
#11 Calculate Parameters
#12 Frequency Sweep ?
Yes
No
Done
[15]
Results Solution Data Available
#13 Statistic Solution Data (profile & convergence tables)
#14 Field Plots (H, B, E…)
#15 Report Plots (H, B, E…)
#16 Calculator
#17 Determine Equivalent Electrical Parameter L & R [equations (4.64) & (4.66)] (not a Maxwell-constituent)

**Figure 4.14** Maxwell flowchart, including solver tasks (gray range) and action items of the user.



Material Name: ferriteK2006_tempdepend_at_100degC

Material Coordinate System Type: Cartesian

| Name | Type | Value | Units | Thermal Modifier |
|---|---|---|---|---|
| Relative Permittivity | Simple | 12 | | None |
| Relative Permeability | Simple | 1700 | | if(Temp > 1000, 19561, if(Temp < -273.15, -5902, 1 + 20 * (Temp - 22))) |
| Bulk Conductivity | Simple | 1 | siemens/m | None |
| Dielectric Loss Tangent | Simple | 0 | | None |
| Magnetic Loss Tangent | Simple | 0 | | None |
| Electric Coercivity | Vector | | | |
| · Magnitude | Vector Mag | 0 | | None |
| Magnetic Coercivity | Vector | | | |
| · Magnitude | Vector Mag | 0 | A_per_meter | None |
| Thermal Conductivity | Simple | 4 | W/m-C | None |
| Magnetic Saturation | Simple | 0 | tesla | None |
| Lande G Factor | Simple | 2 | | None |
| Delta H | Simple | 0 | A_per_meter | None |
| · Measured Frequency | Simple | 9.4e+009 | Hz | None |
| Core Loss Model | | None | w/m^3 | |
| Mass Density | Simple | 4600 | kg/m^3 | None |
| Composition | | Solid | | |
| Specific Heat | Simple | 750 | J/kg-C | None |
| Young's Modulus | Simple | 119000000000 | N/m^2 | None |
| Poisson's Ratio | Simple | 0 | | None |
| Thermal Expansion Coefficient | Simple | 1e-005 | 1/C | None |

View/Edit Material for:
○ Active Design
○ This Product
● All Products

View/Edit Modifier for:
☑ Thermal Modifier

**Figure 4.15** Maxwell dialog box for material property display and editing. Example of user data base extension: Ferrite K2006 with added temperature dependence of the relative permeability.

energy error and is expressed in percentage. Maxwell reports this number after each solution pass and uses it as a measure of solution convergence with respect to the adaptively refined mesh.

At this point, all project features are specified and, after a consistence check (optional), the Maxwell solver can be started (step #5). With this start, the flowchart enters into the machine-processing phase (gray area in Figure 4.14).

The computer autonomously executes the processing steps [15].

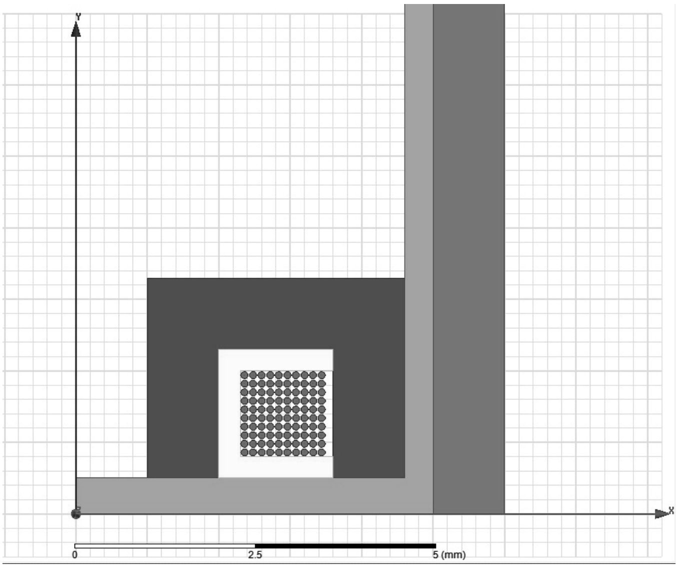


**Figure 4.16** Geometrical model of the Maxwell field simulator: 2D model/right half of the rotationally symmetrical exemplifying ISE in vertical position (see the parts list in Example 4.4). For dimensions, see the scale in the window.

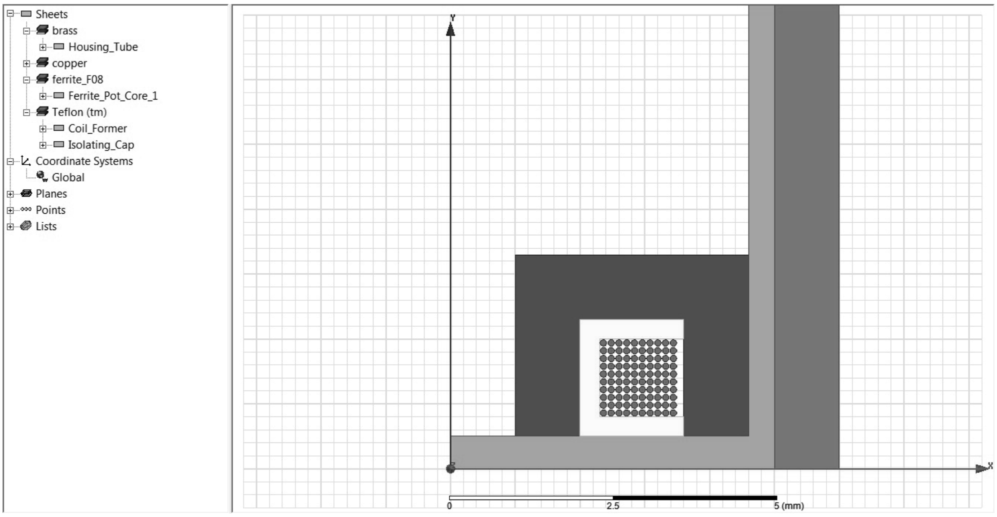


**Figure 4.17** Maxwell GUI detail with the main windows: project tree windows (left) and modeler window with the *rz* geometrical model (see Figure 4.16).

After every main activity, which consists of FE-mashing and field solving (step #6), the software checks if the convergence is fulfilled (step #7). If not, the mesh is refined (step #8) using adaptive mesh refinement techniques and step #6 is reloaded.

If the accuracy of the convergent solution is not achieved (step #9), steps #6 and #7 are repeated as long as the specified number of passes is not exceeded. Contrarily, when the accuracy is achieved or the number of passes exceeds the setup values, the field calculation is stopped (step #10).

If the parameter calculation (step #11) is checked and the software carries on with the three assigned parameters: forces, torques, and inductance/resistance matrix.

In case of an enabled frequency sweep (see step #4), the loop consisting of steps #6 to step #12 is repeated at any sweep frequency $f_S$ in the specified range:

$$f_S = f_{\min} + n \cdot \Delta f, \qquad f_S \in \left[ f_{\min}, f_{\max} \right] \tag{4.75}$$

where $f_{\min}$ and $f_{\max}$ are the specified limits of the sweep range and $\Delta f$ is the sweep step and these were specified at step #4.

From now on, the processing is finished and the data is ready to be evaluated by the user in the postprocessing project steps. Maxwell has a powerful data management and plotting capabilities. Four great facilities are at user's disposal:

1. *Solution data* contains the entire quantity of information related to the executed simulation project. By accessing this option (step #13), the following main statistical reports are also available.
   - *Profile table*, which is a table containing the logbook of performed tasks (real time, CPU time, physical memory used, etc.).
   - *Convergence table* or *plot*, which reports the adaptive convergence information.

- *Mesh statistic table*, which reports information related to FEM mesh in every object.

2. *Field plots* are the representation of basic field quantities as well as of the derived values (losses, densities, etc.) on surfaces or objects (step #14). They can be contour plots (e.g., as in Figure 3.13), vector plots (e.g., as in Figure 3.14), or spectral representations (e.g., as in Figure 3.22).

   The classical plot remains the magnetic flux lines representation. Figures 4.18 and 4.19 illustrate them for the accompanying ISE Example 4.4. The

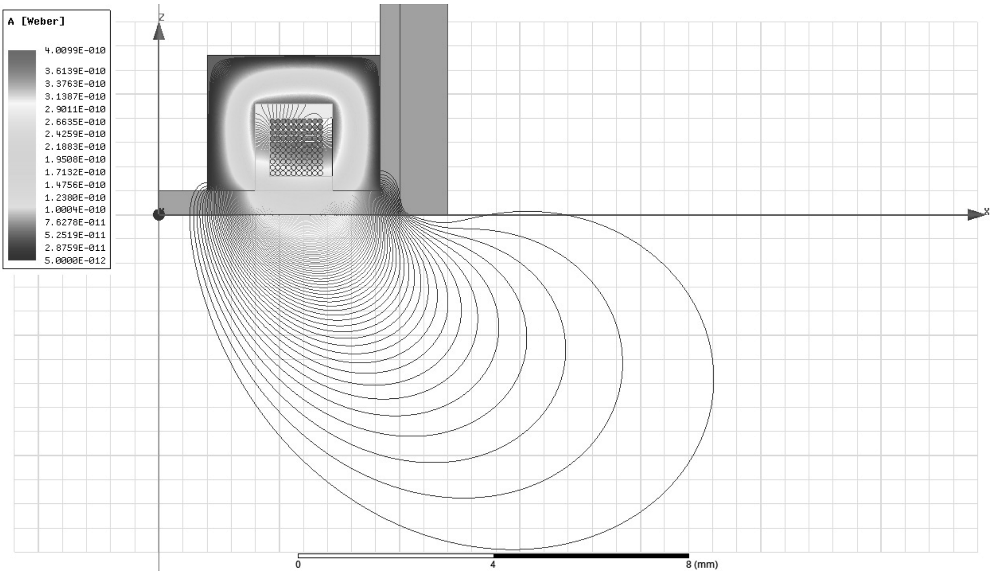


**Figure 4.18** Magnetic field lines of the ISE in Example 4.4 for no target in front of the sensing face.

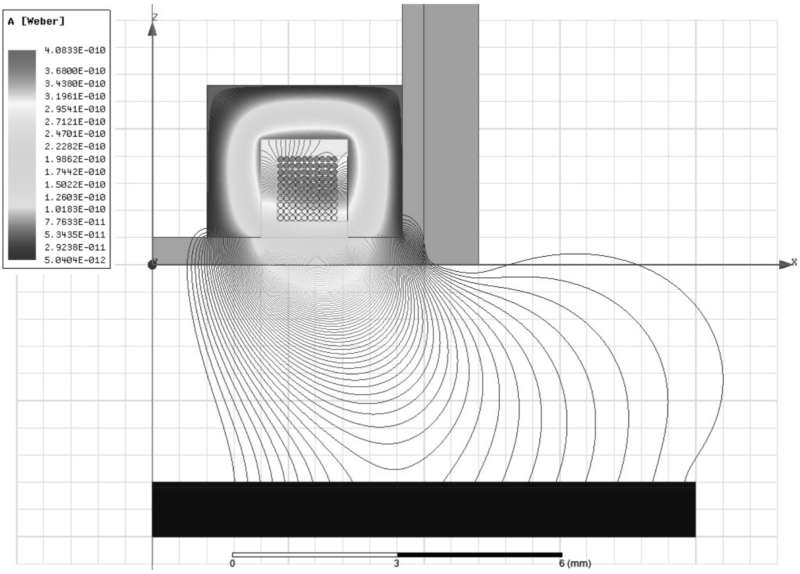


**Figure 4.19** Magnetic flux lines of the ISE in Example 4.4 with steel target (radius of 20 mm and thickness of 2 mm) placed at a distance $d = 4$ mm parallel to the sensor active face.

colors of the field lines (in a color representation) mirror the magnitude of the vector **A** (see (4.53)). The affinity between magnetic field lines and the ferromagnetic target are in evidence.

The attributes of the plots (scale, color map, number of lines, etc.) can be flexibly optimized to better analyze the phenomena or to better view the results.

Applying (3.34) to our example with steel target ($\mu_r = 4000$ and $\sigma = 1.03E7$ S/m), a skin region of about 17.4 $\mu$m results. Skin effects in the range of a few micrometers can still visualized in the Maxwell spectral plot of the flux density $|\mathbf{B}|$ in the proximity target skin (see Figure 4.20). The flux density at the target surface is 6.8 mT.

3. *Report plots* can be generated (step #15) to analyze change in any quantities along a user's defined curve or with respect to any input parameters or time. Rectangular plots multiple traces in single plot area or stacked in more areas as well as in data tabular form can be generated.

   A simple rectangular plot giving the variation of magnitude $|\mathbf{H}|$ of the magnetic field strength along an imaginary measuring line, which has a constant distance of $d = 2$ mm to the sensing face of the ISE, is shown in Figure 4.21.

   The plot allows drawing important conclusions. For example, the magnitude is below 12 A/m and to have an effective full target influence, the target radius should measure 10 to 12 mm (corresponding to the limit of field distribution). Obviously, the distance $d$ can be swept to obtain information about the field intensity reduction with increased distance to the sensor active face.

4. The *field calculator* is a powerful Maxwell tool and needs a special handling because it enables users to postprocess expressions using field quantities and

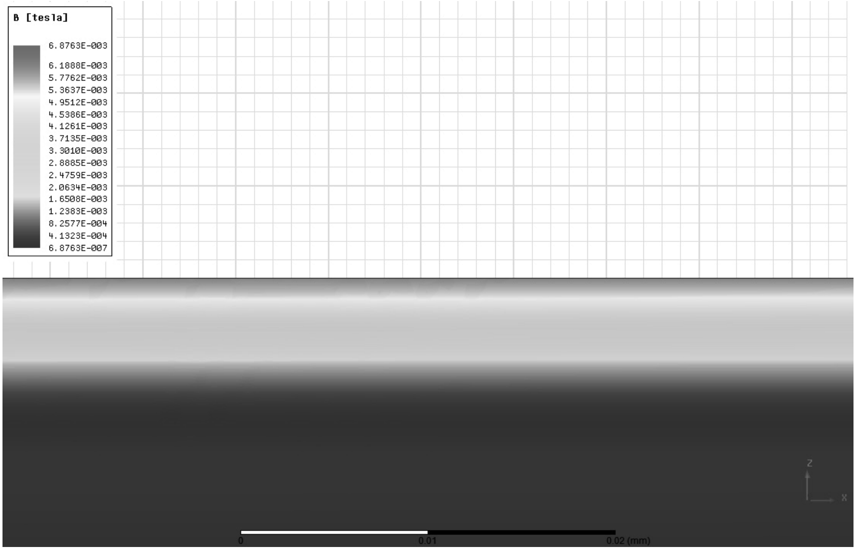


**Figure 4.20** Example 4.4: Skin effect in a transversal section of the steel target (scale 1 $\mu$m per division). A skin region (see (4.60)) in the range of a few micrometers can be measured.

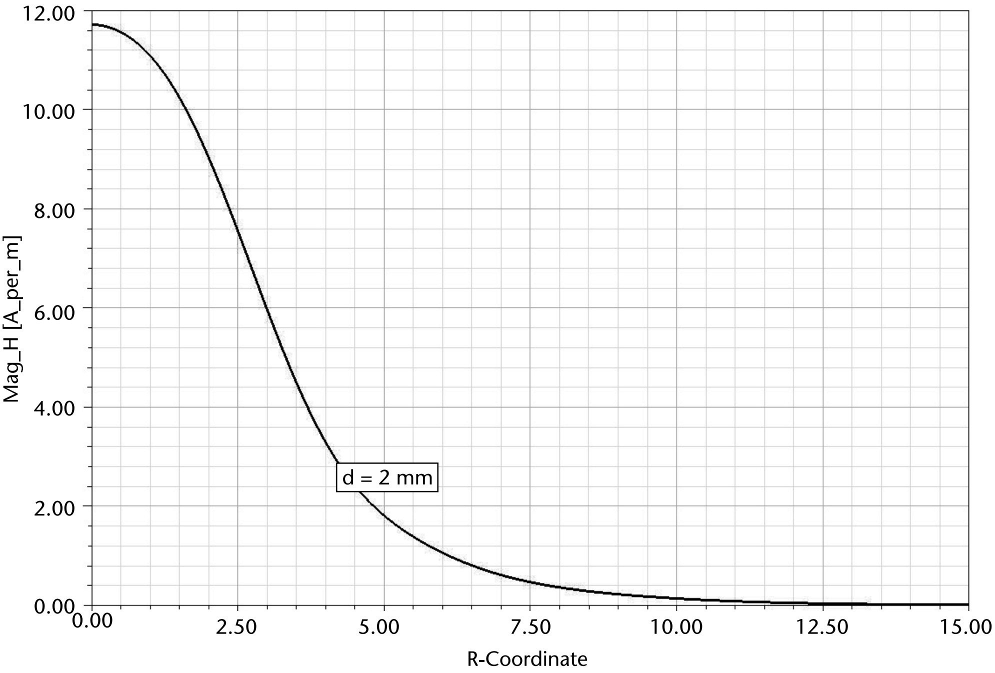


**Figure 4.21** Example 4.4: Field strength |**H**| along a parallel line to the sensing face at d = 2 mm.

derivatives (step #16 and 17). It can perform scalar and vector algebra and calculus operations. It can operate with geometric entities for three purposes:

- Plot field quantities or their derivatives onto geometric entities;
- Perform integration (line, surface, volume) of quantities over specified geometric entities;
- Export field results into a buffer memory stack.

The screen print in Figure 4.22 shows the fields of the Maxwell calculator GUI:

- *Named expressions* field (upper left corner): The user can edit common expressions and store them for later recalls.
- *Solution context* field (upper right corner): If the project contains many solutions, the desired solution has to be checked.
- *Calculator buttons* (underpart), with five columns of buttons: input and output actions as well as general, scalar and vector operations, respectively (see Figure 4.22). The data locations for enter, output or operations are the stack registers.
- *Stack window, register, and commands* (middle range): The calculator interface is structured so that it contains a stack, which holds the quantity of interest in stack register. Stack commands (push, pop, clear, etc.) allow the user to manipulate the stack registers.

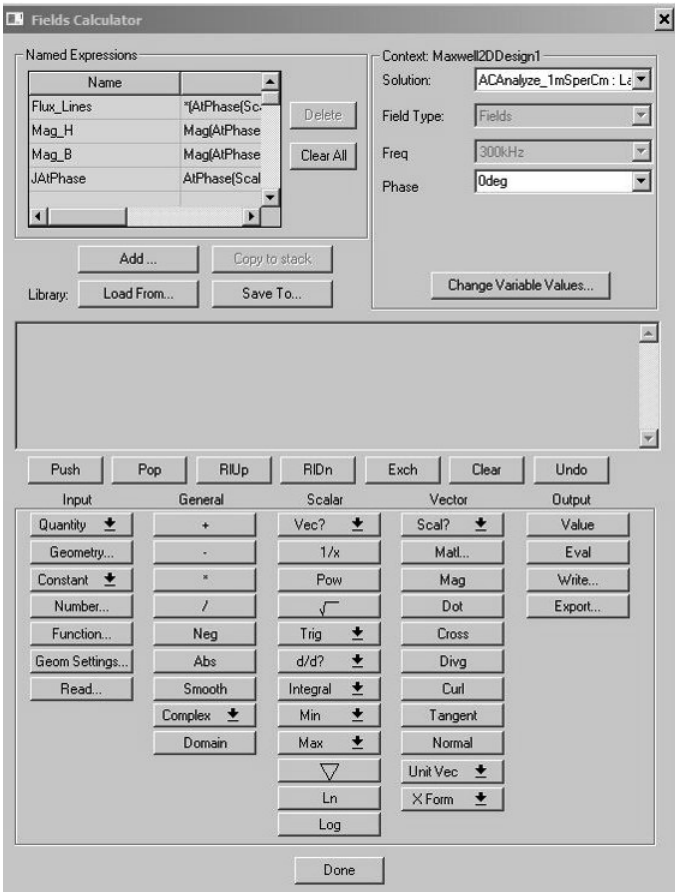
By means of the Maxwell calculator, it is very comfortable to calculate the volume integrals (4.64) and (4.66) for every system entity and to insert these integral numerical values in the final formula below [8, 17], to calculate the Q-factor of the already simulated inductive sensing element:

$$Q_L = \frac{\omega \cdot L}{R_S} = \frac{\omega \cdot \sum_{i=1}^{n} \mu_i \cdot \iiint \mathbf{H}(\mathbf{p}) \cdot \mathbf{H}^*(\mathbf{p})\, dv}{\sum_{i=1}^{n} \frac{1}{\sigma_i} \cdot \iiint \mathbf{J}(\mathbf{p}) \cdot \mathbf{J}^*(\mathbf{p})\, dv} \tag{4.76}$$

where **H*** is the complex-conjugated vector of the magnetic field strength **H** (see (4.64)), **J** and **J*** are the electric current densities: vector and complex-conjugated vector, respectively, $\mu_i$ and $\sigma_i$ are the permeability and conductivity of the entity having the serial number *i*, and $\omega$ is the angular frequency (see (4.5)).

The integrals in (4.76) are computed using the option "*Integral RZ*." This performs integration considering 360° of the geometry around vertical *z*-axis (see Figure 4.13). Thus, the integration is done over the volume of *rz* geometry.

This calculus operation is not included in the Maxwell calculator. We use an additional Microsoft Excel calculation sheet. The numerical values of integrals are


Fields Calculator

Named Expressions

| Name | |
|---|---|
| Flux_Lines | *(AtPhase(Sc |
| Mag_H | Mag(AtPhase |
| Mag_B | Mag(AtPhase |
| JAtPhase | AtPhase(Scal |

Context: Maxwell2DDesign1
Solution: ACAnalyze_1mSperCm : Lɛ
Field Type: Fields
Freq 300kHz
Phase 0deg

Delete
Clear All
Add ...
Copy to stack
Library:
Load From...
Save To...
Change Variable Values...
Push
Pop
RlUp
RlDn
Exch
Clear
Undo
Input
General
Scalar
Vector
Output
Quantity
Geometry...
Constant
Number...
Function...
Geom Settings...
Read...
+
-
*
/
Neg
Abs
Smooth
Complex
Domain
Vec?
1/x
Pow
√
Trig
d/d?
Integral
Min
Max
∇
Ln
Log
Scal?
Matl...
Mag
Dot
Cross
Divg
Curl
Tangent
Normal
Unit Vec
X Form
Value
Eval
Write...
Export...
Done

**Figure 4.22** Maxwell calculator GUI.

transferred in this sheet (via copy and paste), and the remaining algebraic operations (summing, multiplying, and dividing) are automatically performed by Excel software.

As a result of this procedure, the calculated electrical parameters (see Figure 4.1) after simulation of the exemplifying sensing element in Example 4.4 under undamped conditions (no target in front of the active face or $d = \infty$) are:

$$R_{S\infty} = 16.36\Omega,\ \ L_{\infty} = 223\ \mu\text{H and } Q_{L\infty} = 42.8 \text{ at } 500 \text{ kHz} \tag{4.77}$$

## References

[1] Dellsprenger, F., *Smith Software, V4.0*, Bern University of Applied Sciences, 2017, www.fritz.dellsprenger.net.

[2] Dorf, R. C. (ed.), *The Electrical Engineering Handbook*, Boca Raton, FL: CRC Press / IEEE Press, 1997.

[3] Kurz, G., *Oszillatoren*, Heidelberg, Germany: Hüthig-Verlag 1994.

[4] Texas Instruments, *Application Report SNOA936: Measuring RP of an L-C Sensor for Inductive Sensing*, October, 2015.

[5] Webster, J. G. (ed.), *The Measurement, Instrumentation, and Sensors Handbook*, Boca Raton, FL: CRC Press / IEEE Press, 1999.

[6] Griffiths, D. J., *Introduction to Electrodynamics*, Third Edition, Upper Saddle River, NJ: Prentice-Hall, Inc., 1999.

[7] Jagiella, M., S. Fericean, and A. Dorneich, "Progress and Recent Realizations of Miniaturized Inductive Proximity Sensors for Automation," *Sensors Journal*, Vol. 6, No. 6, IEEE, 2006.

[8] Fericean, S., and R. Droxler, "New Noncontacting Inductive Analog Proximity and Inductive Linear Displacement Sensors for Industrial Applications," *Sensors Journal*, Vol. 7, No. 11, IEEE, 2007.

[9] Songoro, H., M. Vogel, and Z. Cendes, "Keeping Time with Maxwell's Equations," *IEEE Microwave Magazine*, IEEE, 2010.

[10] Opera Reference Manual, Oxford, UK, February, 2004, www.rcnp.osaka-u.ac.jp/~sakemi/OPERA/ref-3d.pdf.

[11] *Electromagnetic Sensor Design: Key Considerations when Selecting CAE Software*, white paper, Manitoba, Canada, www.integratedsoft.com.

[12] *Comparison of BEM with FEM*, web publication, Manitoba, Canada, www.integratedsoft.com.

[13] CST Computer Simulation Technology, Darmstadt, Germany, www.cst.com/products.

[14] ANSYS, Canonsburg, Pennsylvania, www.ansys.com/about-ansys.

[15] ANSYS, *Maxwell 17.1 Online Help*, *release 17.1*, Ansys Inc., Canonsburg, PA, April 2016.

[16] Humphries, S., *Finite-Element Methods for Electromagnetics*, eBook, Albuquerque, NM: Filed Precision LLC, 2010.

[17] Hering, E., and G. Schönfelder (Hrsg.), *Sensoren in Wissenschaft und Technik—Funktionsweise und Einsatzgebiete*, 2nd Edition, Wiesbaden, Germany: Vieweg+Teubner Verlag / Springer Fachmedien GmbH, 2018.

CHAPTER 5

# Inductive Sensing Elements: Practical Implementations

After the presentation in Chapter 4 of the analytical, experimental, and numerical computer-assisted techniques to evaluate inductors, which belong to ISEs, the next step in this chapter is to describe how ISEs are made (geometry, materials, and fabrication technologies) and how they work.

The first classification of the inductors contains two main types:

1. Coils (inductors) without core (also called coils with air core);
2. Coils with magnetic core.

A second classification criterion refers to the inductor geometry, and defines the following versions with applications in ISEs:

- Helical round wired coils (single-layer, self-supporting);
- Single-layer/ multilayer cylindrical coils on core former (with or without core);
- Planar flat coils on carrier dielectric material/ insulator (coil on substrate);
- Coil on silicon (coil on chip).

The first five sections of this chapter deal with these specific kinds of inductors that can be used as ISEs. Finally, Section 5.6 focuses on gyrators, which are electronic circuits that cannot be directly implemented in ISEs. However, the circuit can electrically simulate the behavior of an ISE, and thus it is very helpful during real experiments and tests, before the trial and error tuning (coil fabrication and probing) is applied.

## 5.1 Fundamental Inductors: Solenoid and Toroid

Some coils are described in several physical and technical publications as having a primary purpose to illustrate the theoretical knowledge of inductors. They have a lower importance for ISEs. Nevertheless, for information and for the comprehensive character of this volume, this section gives a survey of the most significant fundamental inductors.

### 5.1.1 The Solenoid

The helical solenoid can be considered as the oldest coil of all. It was a test object for a large number of experiments in physics and electrodynamics. According to the Cambridge Dictionary, a solenoid "is a device, consisting of a wire wrapped in the shape of a cylinder that acts like a magnet when electricity goes through it."

It is a self-supporting arrangement with or without a magnetic core, having a high ratio of length to its diameter. It is made from massive winding wire and normally has a winding step greater than the wire diameter (see Figure 5.1).

The magnetic field is concentrated into a nearly uniform field in the center of the long solenoid. The field outside is weak and divergent.

The theory of electrodynamics allows approximating the solenoid inductance. The field density inside the device is approximately constant. Its magnitude $B$, as well as its field strength $H$, is expressed by:

$$B \approx \mu \cdot i \cdot \frac{n}{l} \text{ and } H \approx i \cdot \frac{n}{l} \tag{5.1}$$

where $\mu$ is the magnetic permeability of the solenoid core, $n$ is the number of turns, $l$ is the length of solenoid (in meters), and $i$ is the current through winding.

Substituting the expression of $H$ into (3.3), the inductance $L$ of the solenoid results to be dependent on the magnetic and geometric properties but independent of the current through:

$$L \approx n^2 \cdot \mu \cdot \frac{A}{l} \text{ [henry]} \tag{5.2}$$

where $A$ is the cross-sectional area of solenoid (in meters squared).

This calculation makes use of the long solenoid approximation. It will not give good values for small coils, where the approximation is of limited suitability.

Small inductors for electronics use may be made with an air core (see Section 5.2). For larger values of inductance, magnetic materials are used as core material (see Section 5.3).

Solenoids are adequate for high working frequencies because they have a low parasitic capacitance (≤ 50 pF) and thus a high self-resonance frequency (≥10 MHz). They are suitable for high resonance voltages of series resonant circuits and can easily be tuned by changing their length or slipping the core inside.

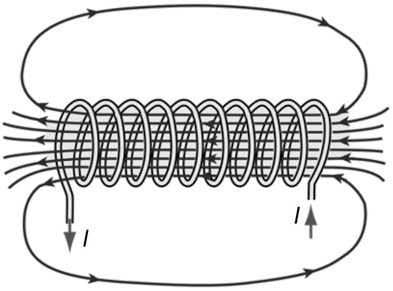


**Figure 5.1** The solenoid and its magnetic field lines.

### 5.1.2 The Toroid

The toroid is by definition a slim-line ring core coil having a uniform winding along its circumference (see Figure 5.2).

The magnetic field is strongly concentrated in the ring-shaped magnetic core and is nearly uniform. The field outside is negligible.

The theory of electrodynamics allows approximating the solenoid inductance in this case as well. Finding the magnet field inside the toroid is a good example of the power of Ampere's law:

$$\oint_{DL} B \cdot dl \approx \mu \cdot I_{NC_S} \tag{5.3}$$

The meaning of this closed-line integral around a closed-curve is that the current enclosed by the dashed line $DL$ is just the number of the winding loops multiplied by the current in each loop ($I_{NCS}$ is the net current passing through the total surface $S$):

$$B \cdot 2\pi r \approx \mu \cdot n \cdot i \text{ thus } H \approx i \cdot \frac{n}{2\pi r} \tag{5.4}$$

where $r$ is toroid radius to center line (in meters) and $i$ is the current through winding.

Substituting now this expression of $H$ into (3.3), the inductance $L$ of the toroid results to be dependent on magnetic and geometric properties but independent of the current through:

$$L \approx n^2 \cdot \mu \cdot \frac{A}{2\pi r} \,[\text{henry}] \tag{5.5}$$

where $A$ is the cross-sectional area of the coil (in meters squared).

The analogy between (5.2) and (5.5) is obvious; the difference is just given by the effective length of the core.

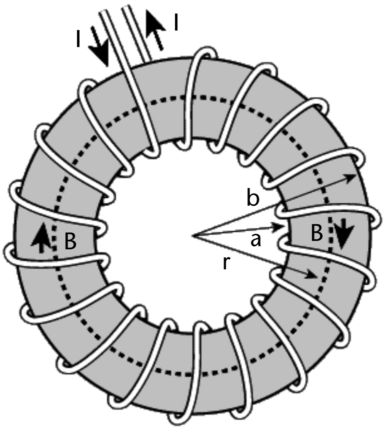


**Figure 5.2** The toroid and the flux density vector **B**.

## 5.2 Wire-Wound Coils with Air Cores

The calculation of the inductance of a classical cylindrical coil is nowadays an easy exercise. The internet provides several $L$ calculators. The user has to give the coil features (number of turns and geometrical dimensions) and an empirical formula, which is stored in the software which then computes and immediately displays the result. The tolls generally go back to the classical formula of Wheeler [1], which has become a basic reference on this field.

For information and for the comprehensive character of this volume, Table 5.1 shows relevant calculation formulas for measuring units of the British/American system and for the International System of Units, respectively [2, 3]. All of the following considerations are related to winding coils without core.

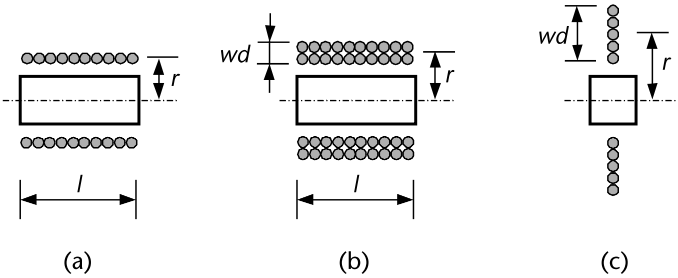
*Example 5.1* A short coil having $n = 18 \cdot 8 = 144$ turns, $r = 4.1$ mm (0.1614 in), $l = 5.3$ mm (0.2086 in) and $wd = 2.2$ mm (0.0866 in) is represented in Figure 5.4. The inductance calculated with (5.7) is $L = 116.38\ \mu$H.

Certainly, better accuracy and more information can be obtained using CAE-tolls, as shown in the previous chapters. As result of a Maxwell simulation, the Maxwell calculator (see Figure 4.22) provided, using (4.64) and (4.66), a computed inductance of $L = 118.5\ \mu$H and a loss resistance of $R_S = 2.14\Omega$ (at 1 kHz). Because

**Table 5.1** Calculation Table of the Cylindrical Coil Inductance

| | | *Inductance* | |
|---|---|---|---|
| *Coil Type* | *Pictogram in* | *These Formulas Require Units in Inches* | *This Formula Requires Units in Millimeters* |
| Single-layer | Figure 5.3(a) | $L = \frac{n^2 \cdot r^2}{9 \cdot r + 10 \cdot l}$ [$\mu$H] (5.6) | $L = \frac{n^2 \cdot (d/10)^2}{4.5 \cdot d + 10 \cdot l}$ [H] (5.9)<br>$(\text{error } < 1\% \text{ if } l \geq 0.5 \cdot d)$ |
| Multilayer, multirow | Figure 5.3(b) | $L = \frac{0.8 \cdot (n^2 \cdot r^2)}{6 \cdot r + 9 \cdot l + 10 \cdot wd}$ [$\mu$H] (5.7) | |
| Multilayer, single-row | Figure 5.3(c) | $L = \frac{n^2 \cdot r^2}{8 \cdot r + 11 \cdot wd}$ [$\mu$H] (5.8) | |

$n$ = number of turns; $r$ = coil radius; $d$ = coil diameter; $l$ = coil length; $wd$ = winding depth.



(a) (b) (c)

**Figure 5.3** Pictograms of the inductors discussed in Table 5.1

of high investment costs with the CAE tolls acquisition, the aforementioned formulae keep their practicable application.

The inductance of a coil depends on its geometrical characteristics, the number of turns and the method for winding the coil. The larger the diameter and the larger number of winding turns, the greater its inductance. If the coil is tightly wound, turn to turn, then it will have more inductance than a not tightly wound coil, with gaps between the turns.

## 5.3 Wire-Wound Coils with Magnetic Cores

The vast majority of inductors for ISEs are coils with magnetic cores. A core theoretically provides a so-called useful field amplification effect.

The presence of this core is a reason for the following positive actions:

- *Focusing and collimating* of the flux lines. The core concentrates the field lines, carries them, and prevents parasitic field dispersions. A very representative example of this functionality is shown in Figure 3.9(a).
- *Orientation and guidance* of the flux lines towards the target that moves (see Figure 3.22). This effect considerably increases the system sensitivity.
- *Shielding* from environment's parasitic influences. The core reduces the electromagnetic disturbing actions of vicinal metallic objects, making embeddable IS applications possible (see Figure 2.2(a)).

The cores generally consist of the one of following magnetic materials:

- Ferrites;
- Permalloy, Mu-metal, or other magnetic alloys;
- Iron and iron alloys (iron with silicon, cobalt, etc.).

Chapter 6 comprehensively deals with these magnetic materials for cores of ISEs.

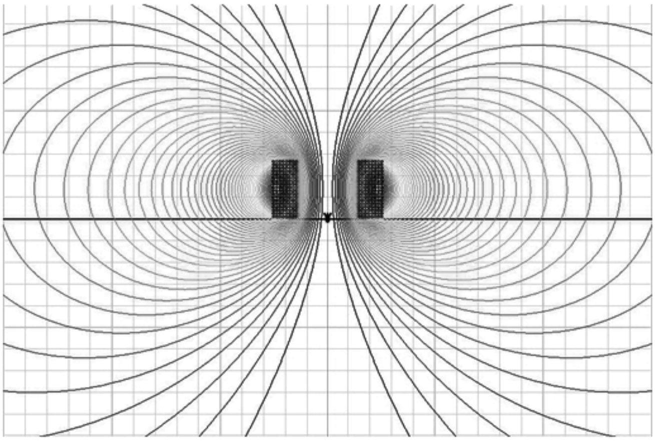


**Figure 5.4** Magnetic field lines of short coil with an air core (transversial section view), as considered in Example 5.1.

### 5.3.1 Inductance of Wire-Wound Coils with Magnetic Cores

Elaboration of analytical expressions for the ISE inductance and its variation as result of a target or core travel is generally a difficult task (see Section 3.3). Equation (5.2) could theoretically be used.

However, for simple coil geometries with magnetic coils, which have shapes like those in Figure 5.3, there are calculation formulas in technical publications, or found on the internet, similar to (5.6) through (5.9).

When the length of a short inductor is higher than its radius (see Figure 5.3(a)), the inductance can be calculated by using the formula (SI measuring units):

$$L = \frac{\mu_0 \cdot \mu_r \cdot n^2 \cdot \pi r^2}{l + 0.9r} \tag{5.10}$$

where $\mu_0$ is the absolute permeability of free space ($4\pi \cdot 10^{-7}$ H/m), $\mu_r$ is the coil core relative permeability, $n$ is the number of coil turns; $l$ and $r$ are the coil length and radius, respectively.

For the reverse situation $l < r$, (5.10) needs a small adjustment to maintain accuracy:

$$L = \frac{\mu_0 \cdot \mu_r \cdot n^2 \cdot \pi r^2}{1.1l + 0.8r} \tag{5.11}$$

As presented in Chapter 6, the permeability of the coil core has more dependences, the most significant being the temperature dependence and the disaccommodation (see (6.29)).

If the temperature changes from $T_1$ to $T_2$, and the permeability correspondingly varies from $\mu_{r1}$ to $\mu_{r2}$, the temperature drift of the inductance, calculated with (5.10) or (5.11), results from the formula:

$$\frac{L_2(T_2) - L_1(T_1)}{L_1(T_1)} = \frac{\mu_{r2} - \mu_{r1}}{\mu_{r2} \cdot \mu_{r1}} \cdot \mu_e \tag{5.12}$$

where $\mu_e$ is the effective permeability (see (6.2)), which is given for each core in the individual data sheet.

Similarly, a change in inductance over a period of time at a constant temperature can be calculated with the aid of disaccommodation factor $D_F$ (see (6.29)):

$$\frac{L_1(t_1) - L_2(t_2)}{L_1(t_1)} = \mu_e \cdot D_F \cdot \log\frac{t_2}{t_1} \tag{5.13}$$

where $L_1(t_1)$ and $L_2(t_2)$ are values at two time intervals $t_1$ and $t_2$ after a strong disturbance.

The core coils generally have much more complex geometry and such formulas are not applicable. A practical method for technical core types consists of the applying of the core factor.

### 5.3.2 Core Factor and Effective Core Parameters

For the calculation of an uneven magnetic core, and of the resulting coil made with this core, a method for convenience can be used according to IEC 60205. This method is based on so-called effective dimensions: length $l_e$, surface $A_e$, and volume $V_e$.

These dimensions define a hypothetical equivalent ring core, which has the same magnetic properties with the uneven soft core. In thin ring cores (internal diameter greater than 0.8 times the external diameter), a constant magnetic flux can be expected. Thus, the ideal ring core is characterized by an equivalent reluctance:

$$R_{m\_e} = \frac{1}{\mu_i \cdot \mu_0} \cdot \frac{l_e}{A_e} \tag{5.14}$$

The reluctance of the uneven core can be expressed by:

$$R_{m\_uec} = \frac{1}{\mu_e \cdot \mu_0} \cdot \sum \frac{l_e}{A_e} \tag{5.15}$$

The condition of the thin ring core is not met for technical core types. Form factors $C_1$ and $C_2$ are introduced to be able to apply the formulae derived so far for ring cores to other core types:

$$C_1 = \sum_{i=1}^{m} \frac{l_i}{A_i} \quad \left[\text{mm}^{-1}\right] \text{ and} \tag{5.16}$$

$$C_2 = \sum_{i=1}^{m} \frac{l_i}{A_i^2} \quad \left[\text{mm}^{-3}\right] \tag{5.17}$$

where $m$ is the number of core segments of a constant effective magnetic cross section, $l_i$ is the magnetic path length of the segment $i$, and $A_i$ is the magnetic effective cross section of the path length $l_i$.

It is important to note the exception that the ferrite core providers use the millimeter unit instead of the SI unit (meter) for these parameters.

Moreover, the core factors can be used to determine the effective core parameters of surface $A_e$, length $l_e$, and volume $V_e$, which are defined as follows:

$$A_e = \frac{C_1}{C_2} \quad \left[\text{mm}^2\right] \tag{5.18}$$

$$l_e = \frac{C_1^2}{C_2} \quad \left[\text{mm}\right] \tag{5.19}$$

$$V_e = \frac{C_1^3}{C_2^2} = A_e \cdot l_e \left[\text{mm}^3\right] \tag{5.20}$$

The core providers usually state the core factors $C_1$ and $C_2$ for the offered cores. In addition, they could specify the three effective core parameters.

Having these values, the user can estimate the inductance $L$ of a coil made with this technical core type:

$$L = \frac{\mu_e \cdot \mu_0 \cdot n^2}{C_1} \tag{5.21}$$

where $n$ is the number of turns.

In addition, the field main quantities can be estimated. If the effective coil current $I$ is known, then the peak value of the magnetic field strength is:

$$H_{\text{peak}} = \frac{n \cdot I \cdot \sqrt{2}}{l_e} \tag{5.22}$$

The effective cross section $A_e$, the effective value of the voltage $V$, and the frequency $f$ of the AC current can be used to obtain the peak value of flux density:

$$B_{\text{peak}} = \frac{V \cdot \sqrt{2}}{2\pi f \cdot n \cdot A_e} \tag{5.23}$$

We intentionally called such calculations estimations because they are based on approximations and on typical values.

*Example 5.2* To give a feeling of accuracy, the ISE used in Example 4.4 is now resumed. Its inductance will be calculated using the specified effective core parameters and the result will be compared with the inductance value provided by a precise field simulation made with the Maxwell solver.

The coil of the ISE schematically represented in Figure 3.20 and accurately modeled in Figure 4.16 uses a technical ferrite pot core Sch9 F08, manufactured by Neosid/Germany (type: Sc, material: F08) (see Table 6.1). The core has an external diameter $\varphi_d = 9.2$ mm and a height $h = 2.8$ mm (see the fourth example in the *Pot Cores* row in Table 6.3).

For an arrangement of two such cores placed face-to-face (closed magnetic loop), the Neosid data book specifies [4]:

- $C_1 = 1.25 \text{ mm}^{-1}$;
- $l_e = 12.5$ mm;
- $A_e = 10 \text{ mm}^2$;
- $V_e = 125 \text{ mm}^3$.

For the initially considered number of turns $n = 100$, (5.21) gives a result of $L \approx 7030\ \mu\text{H}$, which is quite far from the value $L \approx 350\ \mu\text{H}$ provided by the Maxwell field simulator for a similar situation.

### 5.3.3 Losses Caused by Cores: Total Formula of the Inductor Impedance at Full Length

In every magnetic circuit, including the ISEs, losses occur. The losses of technical inductors, which represent different ISEs, were mentioned many times over the previous chapters. In this context, (3.24) and (3.35) generally express their impedance by:

$$\underline{Z}(s) = R_S(s) + j\omega \cdot L(s) \tag{5.24}$$

where the variable $s$ is the distance between target and coil and $\omega$ is the angular frequency, which corresponds to the working frequency $f$ ($\omega = 2\pi f$).

The quantities $R_S$ and $L$ are the coil parameters series loss resistance and inductance, and both are influenced by target. As stated throughout the previous chapters (e.g., Section 3.3.3), the loss resistance $R_S$ is essentially the sum of four components, which are results of the aforementioned effects and are independently influenced by the target.

The same consideration, this time for any real inductor is expressed in (4.12). The basic classification of the losses includes three major loss categories:

1. *Copper losses* (a traditional term from classical electrical engineering) are generated by the ohmic resistance of the coil winding. It is the classic joule heating process, by which the passage of the electric current through the electrically conductive winding medium produces heat.
2. *Iron losses* (also a traditional term from the classical electrical engineering) are cyclic magnetization losses in terms of heat development, which are caused by the AC core magnetization, as well as losses caused by the eddy currents that occur in the conductive core material and generate joule heating. Finally, the magnetic aftereffect losses have to be considered, which additionally produce heating. These major losses are characteristic for every technical inductor having a magnetic core.
3. *Additional losses* in the metallic parts of the ISE and in the movable target are optional losses caused by the eddy currents that occur in such metallic parts and obviously generate additional joule heating.

All these losses have an ohmic character and can be analytically expressed for calculating the corresponding loss resistance component. The effects are independent of each other and the electrical equivalent series circuit results applying the superposition theorem. The general physical model of the impedance $\underline{Z}$ of an ISE therefore comprises the inductance $L$ and a series chain of loss resistances (see Figure 5.5(a)).

Due to the total loss resistance $R_S$, the voltage drop $\underline{V}$ across the ISE lags the voltage in the inductive element by angle delta. As a result, the phase angle between the voltage $\underline{V}$ and the current $\underline{I}$ in the ISE is given by:

$$\arg\underline{U} - \arg\underline{I} = 90° - \text{Delta} \tag{5.25}$$

During the sensor operation, all components in Figure 5.5(a) are influenced by the object to be detected. However, the degrees of influence are different, depending on the system type (see Sections 3.3.1–3.3.7), the implemented materials, and

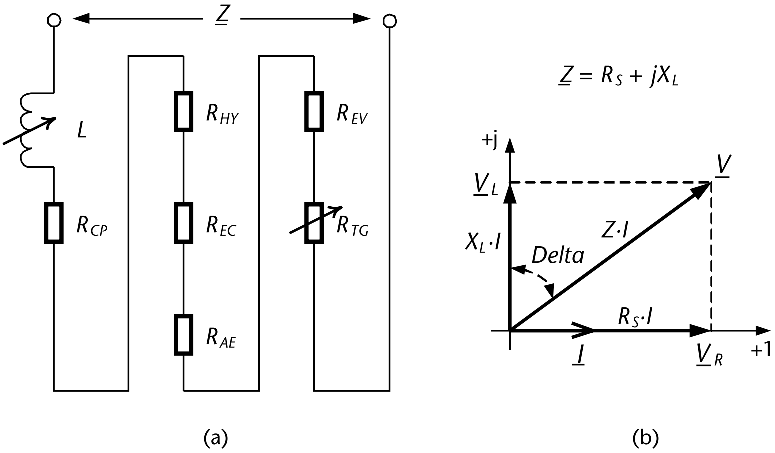


**Figure 5.5** (a) General equivalent electrical circuit of the impedance of an ISE, and (b) phasor diagram of the corresponding sinusoidal waveforms (voltages and current). Reference axis is the current $\underline{I}$ axis (series representation).

the operating conditions. As an example, variable values $L$ and $R_{TG}$ suggest they are strongly influenced by the target at systems based on eddy current evaluation.

### 5.3.3.1 Copper Losses

The copper losses have a generally predominant value and are strongly dependent on the ISE working frequency.

At low frequencies, the skin effect is negligible and the ohmic loss resistance $R_{CP}$ of the coil wire can be expressed in different ways as a DC resistance. The first possibility consists of the classical resistance calculation for a single turn (average value) and of the multiplication with the number $n$ of turns:

$$R_{CP} = n \cdot \frac{\rho \cdot l_N}{A} \tag{5.26}$$

where $\rho$ is the copper resistivity (typically 1.68E-8 Ωm or 1.72E-8 for annealed copper), $l_N$ is the average length of a core turn, and $A$ is the cross section of the coil wire.

A second alternative recommended by some core manufacturers uses the expression of coil DC resistance by means of the resistance factor $A_R$, which is analogous to the $A_L$ (compare with (6.16)):

$$R_{CP} = n^2 \cdot A_R \tag{5.27}$$

where $n$ is the number of turns and the resistance factor $A_R$ is given by:

$$A_R = \frac{\rho \cdot l_N}{F_{CP} \cdot A_N} \tag{5.28}$$

where $\rho$ is the wire resistivity (usually cooper), $l_N$ is the average length of a turn, $A_N$ is the cross section of the winding and $F_{CP}$ is the cooper space factor. For a rough estimation, a copper filling factor $F_{CP} = 0.5$ is sufficient.

The manufacturers offer suitable coil formers and specify the factor $A_R$.

The increase of the working frequency causes the skin depth $\delta$ (see (3.30)) to nonlinearly drop (e.g., for copper $\delta = 0.66$ mm at 10 kHz and $\delta = 0.066$ mm at 1 MHz). At higher frequencies, the skin effect has to be considered unconditionally. If the value of skin depth becomes lower than the radius of the coil wire, the current concentrates and flows in a thin circular range below the wire surface. The reduction of the effective cross section $A$ of the wire leads to a substantial increasing of the losses and—in this way—of the resistance $R_{CP}$ (compare with (5.26)).

#### 5.3.3.2 Iron Losses

*Magnetization Losses*

The magnetization losses are a consequence of the magnetic materials hysteresis, understood as the delay between the resulting flux density and its modifying cause, namely the field strength. This phenomenon is geometrically represented by the hysteresis curve (see Figure 6.1) of the core material and the surface area of this representation is a measure of the hysteresis losses. During every magnetization cycle, a certain amount of energy is consumed from the magnetization current source and heat dissipated in the ferrite core. This process is represented by magnetic losses resistance $R_{HY}$ that is proportional to the working frequency $f$ and to the surface area bordered by the hysteresis curve, which implies permeability $\mu$ and field strength $H$. To give the expression of the loss resistance, quantities $R_{HY}$, $f$, and $\mu$ can formally be linked up by means of a hysteresis coefficient $k_{HY}$:

$$R_{HY} = k_{HY} \cdot f \cdot \mu \tag{5.29}$$

Similar to (6.11) for $R_S$, the component $R_{HY}$ can be expressed by means of the hysteresis loss factor $\tan\delta_h$. After the extraction of the measured tan values $\tan\delta_m$ from the plots shown in Figure 6.3, it is possible to calculate loss resistance $R_{HY}$:

$$\frac{R_{HY}}{\omega L} = \tan\delta_m (B = 0) - \tan\delta_m \left(B_1 = 1.5 \text{ mT}\right) \tag{5.30}$$

ISEs are generally exposed to weak magnetic field strength. For low AC field amplitudes, the shape of the well-known B-H curve (see Figure 6.1) substantially changes approaching the form of an ellipse with its center placed in the origin of the plane B-H and its major axis inclination imposed by the limits of the deflection on the H-axis.

Correspondingly, for applications working with $H < H_C$ (a majority of ISEs), reference [5] proposes to “linearize the dependence B to H” replacing the general hysteresis curve by the so-called Rayleigh cycle (see Figure 5.6), which can be analytically expressed and has a calculable area. The equation of the magnetization elliptic trajectory is:

$$B - B' = \mu_i \cdot (H - H') \pm \frac{a}{2} \cdot (H - H')^2 \tag{5.31}$$

where $(B - B')$ is the flux density change provoked by the field variation $(H - H')$, $\mu_i$ is the initial permeability, and $a$ is a specific material constant. The sign between terms is + or − depending if $H > H'$ or $H < H'$.

The dissipation is proportional to the surface area $A_{HY}$ bordered by the Rayleigh curve and $R_{HY}$ becomes [5]:

$$R_{HY} = k^{*}_{HK} \cdot f \cdot L \cdot B_{M} \tag{5.32}$$

where $k^{*}_{HK}$ is now a material characteristic, dependent on $\mu_i$, $L$ is the coil inductance, and $B_M$ is the elongation of the flux density.

*Eddy Current Losses*

The eddy current losses caused in the core are a result of the AC magnetization. According to Faraday's law of induction, circular current-flow paths, so-called eddy (or Foucault) currents result in the core. They detract from the field active power, which is converted into heat. In this way, heat losses occur. This effect is a function of material permeability $\mu$ and increases with the square of frequency $f$, so that the corresponding losses resistance $R_{EC}$ can be expressed by:

$$R_{EC} = k_{EC} \cdot f^2 \cdot \mu = \frac{k^{*}_{EC}}{\rho} \cdot f^2 \cdot L \tag{5.33}$$

where $k_{EC}$ is the so-called eddy currents losses coefficient, $k^{*}_{EC}$ is a material characteristic and $\rho$ is the core resistivity.

For the cores used in ISE, there is generally no need to reduce the eddy current losses by constructive means. However, the fragmentation of the core into a large number of components (wires, laminates, rolled-up foil, etc.) that are fully insulated to each other is a classical improvement that could be used. This measure increases the total cross area of the current and correspondingly reduces the resistance $R_{EC}$.

*Jordan Magnetic Aftereffect Losses*

The Jordan magnetic aftereffect losses occur due to the fact that a stepwise change of the magnetic field strength cannot be inertia-free followed by the flux density. During periodical magnetization and reversals of magnetization, a phase shift between field and magnetization results. The final consequences are aftereffect losses. The corresponding losses resistance $R_{AE}$ is also proportional to $f$ and $\mu$ and can formally be expressed by a similar equation to (5.29):

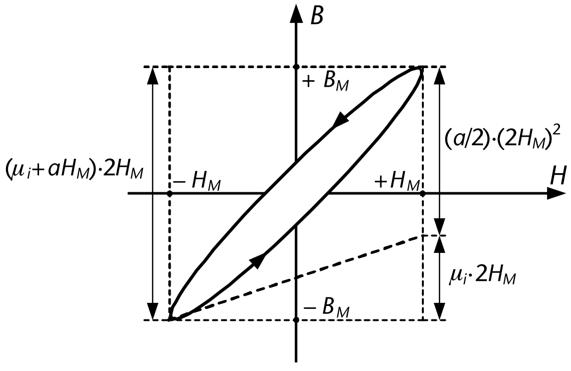


**Figure 5.6** The shape of the Rayleigh hysteresis cycle [5].

$$R_{AE} = k_{AE} \cdot f \cdot \mu = k_{AE}^{*} \cdot f \cdot L \tag{5.34}$$

At this time, it is interesting to observe that every action that influences the core permeability—not only geometrical changes (see Section 3.2 for IS-based geometry interaction) but also magnetoelastic actions (see Section 3.2 for IS-based material interaction)—leads to a change of the inductance but also a variation of the loss resistance. This occurs if the working frequency is high enough. The ever-existing fact can be successfully exploited to increase the sensor sensitivity.

#### 5.3.3.3 Additional Losses

The additional losses are caused by various metallic parts of the ISE (sensor housing, end plates, pushrod, etc.). They are unwanted losses and should be minimized as far as possible. In the equivalent electrical circuit (Figure 5.5), the losses generated by environmental metallic parts in the ISE as well as in the surrounding area of the IS are represented by the losses resistance $R_{EV}$.

##### *5.3.3.4 Losses Caused by Target*

Last, but not least, if the IS performs the detection of a mobile plunger (e.g., in Sections 3.3.4 and 3.3.5) or movable target (e.g., in Section 3.3.6), the losses caused by this metallic object represent sensor input information and should strongly be emphasized. The higher the corresponding losses resistance $R_{TG}$, the easier the sensing operation. Field simulation programs like Maxwell are the best way to determine $R_{EV}$ and $R_{TG}$.

The general expression of the loss resistance $R_S$ (see (5.24)) of the impedance $\underline{Z}$ of an ISE (see Figure 5.5(a)) is finally given according to the Jordan expression:

$$R_S = R_{CP} + \left(R_{HY} + R_{EC} + R_{AE}\right) + \left(R_{EV} + R_{TG}\right) \tag{5.35}$$

For quick, preliminary evaluation of the losses in the core, (5.32) to (5.34) can be brought together in the formula:

$$R_{S\_C} = R_{HY} + R_{EC} + R_{AE} = k_{S\_C} \cdot 2\pi f \cdot L \tag{5.36}$$

where the proportionality factor $k_{S\_C}$ is a function of frequency and flux density in the core and has an order of magnitude of $10^{-5}$ according to [5].

*Example 5.3* To give the reader a concrete understanding about the magnitude order of these losses resistances for a real ISE, Example 4.4 (Section 4.3.4) is resumed here. Example 4.4 uses the ISE of an IPS, model IA-M12. The model was continuously processed by the Maxwell field solver up to the computation of the impedance components without target: $R_{S\infty} = 16.36\Omega$ and $L_\infty = 223\ \mu\text{H}$ (see (4.77)). The present example goes back to the Maxwell calculator and extracts the already available $R_S$-subcomponents (see Table 5.2). Three cases are considered: ISE without target in front, which is equivalent to $s = \infty$ (note the total values above mentioned), with a steel target (ferromagnetic) as well as with an aluminium target (paramagnetic) with identical sizes and at the same distance $s = 2$ mm ahead.

**Table 5.2** Losses Equivalent Resistances of the ISE of an IPS, Model IA-M12

| | | *Working Conditions of the ISE* | | |
|---|---|---|---|---|
| *Part of ISE* | *Losses Resistance (Ω)* | *No Target in front of the ISE* | *Steel Target at s = 2 mm* | *Aluminum Target at s = 2 mm* |
| Winding | $R_{CP}$ | 13.655 | 13.647 | 13.665 |
| Ferrite pot core | $R_{HY} + R_{EC} + R_{AE}$ | 0.00487 | 0.00507 | 0.00458 |
| Metal tube | $R_{EV}$ | 2.704 | 3.501 | 1.546 |
| Target | $R_{TG}$ | 0.000 | 4.156 | 1.338 |
| | $R_S$ (total resistance) | 16.363 | 21.309 | 16.554 |
| | Inductance $L$ [$\mu$H] | 222.92 | 226.21 | 216.04 |

See Example 5.3.

The exemplified ISE belongs to the IS_EC evaluation systems (see Section 3.3.6). The system works at 500 kHz, the skin effect has a predominant influence in the winding wire and leads to high values of $R_{CP}$.

In contrast, the core losses and therefore the sum $R_{HY} + R_{EC} + R_{CM}$ are very low (a few of milliohms) as a result of the low magnetic field values and good core ferrite quality. The diamagnetic brass tube playing the housing protection role (see Figure 3.20) induces a remarkable unwanted loss $R_{EV}$, which essentially reduces the system sensitivity on the target travel (see Section 1.3.3).

Finally, the actuation of these sensor types are based on the evaluation of the losses induced by the target $R_{TG}$ and also of the target influence on the ISE inductance. Obviously, these variations are dependent on the target absence/presence, on the distance between ISE and target (see Figure 3.20) and also on the target magnetic properties (ferromagnetic, dia/paramagnetic, etc.).

## 5.4 Printed Flat Spiral Coils

The flat spiral (planar) coils realized on printed circuit board (PCB) play a role with an ever-expanding importance as well as the presence and extension in ISEs.

The reason for these consists of lower production costs and material needs of new constructive perspectives and opportunities (miniaturization, optimal shapes and positions, etc.), and of better relevant technical features as a low sample deviation, good reproducibility, and low tolerances of coil pairing.

The large majority of PCB coils have the geometry of the classical Archimedean spiral. It is the locus of points representing the locations over time of a point moving away from an origin point with a constant speed along a line, which, at the same time, rotates with constant angular velocity (see Figure 5.7).

This evolution is described in polar coordinates ($r$, $\alpha$) by:

$$r = a + b \cdot \alpha \tag{5.37}$$

where $a$ and $b$ are real numbers. The parameter $a$ defines the spiral excursion (start and end points), while $b$ controls the distance between successive turns (spacing).

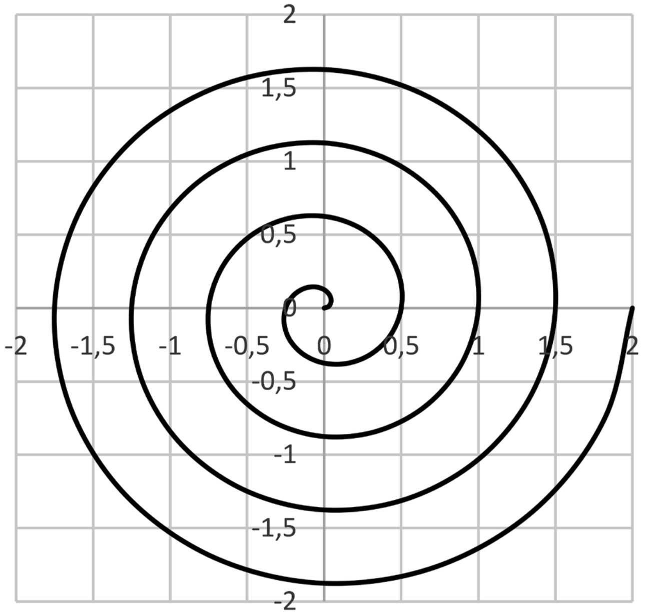


**Figure 5.7** Four $2\pi$ turnings of a branch of an Archimedean spiral.

The expression in polar coordinates can be rewritten in a parametric form in the Cartesian coordinates system (see Figure 5.8):

$$\begin{aligned} x &= r \cdot \cos\alpha = (a + b\alpha) \cdot \cos\alpha \\ y &= r \cdot \sin\alpha = (a + b\alpha) \cdot \sin\alpha \end{aligned} \tag{5.38}$$

and can accordingly be represented in the $xy$-plane (see Figure 5.7).

A property of the Archimedean spiral that makes it reasonable for flat coil implementations is that any ray from the origin intersects successive turnings of the spiral in points with a constant separation distance, hence its second name arithmetic spiral. For the real coil, this fact leads to easier design and fabrication.

The large majority of flat coil models is based on the classical circular spiral. For example, the coil model in Figure 5.9(a) represents a coil having 4 turns and an origin point different from zero ($a_{\text{initial}} = 0.75$, $a_{\text{final}} = 2.75$). Other derived geometries with octagonal turning shape (see Figure 5.9(b)), hexagonal or—at the limit—square spirals are disclosed in technical publication, however less used [6].

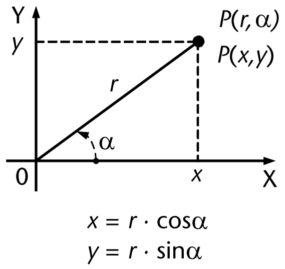


**Figure 5.8** Conversion from polar to Cartesian coordinates.

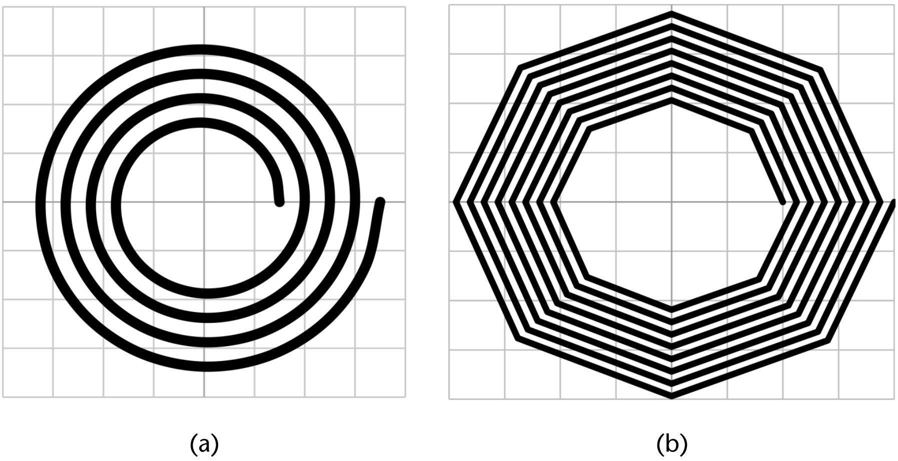

(a) (b)

**Figure 5.9** (a) Example of a round PCB coil geometry with Archimedean spiral shape (4 turnings), and (b) example of an octagonal PCB coil geometry with Archimedean spiral shape (8 turnings).

All these coils are single-layer or multilayer planar inductors and are made on a PCB. The PCB concept mechanically supports and electrically connects electronic parts or electrical components. It purposefully uses conductive tracks, pads, and other features etched from one or more sheet layers of copper, which are laminated onto and/or between sheet layers of a nonconductive substrate. Components are generally soldered onto the PCB to both electrically connect and mechanically fasten them to it.

The procedure to realize simple PCB coils is a standard photolithography process consisting of the following seven steps:

1. A photolithography mask which has the negative image of the planar coil (transparent areas finally generate conductive lines) is made by means of high accuracy laser plotters.
2. A carrier material consisting of epoxy resin or glass film and a single- or double-sided layer covered with laminated copper is used.
3. After a precleaning process of this carrier material, a photosensitive resist (laminar) is dry-film laminated.
4. The mask is applied on the prepared carrier and the ultraviolet-light (UV) exposure is performed.
5. The laminar is developed from carrier surfaces. In the already light-exposed areas, the laminar is polymerized, gets a hard adherence, and remains on the core surface. The temporary image of the hardened resist is the coil image.
6. The nonexposed, uncovered copper is removed by the etching process.
7. Finally, the blue photoresist, which protected the copper image, is dry stripped off and the exact pattern required is available.

By far, the most common material used today for PCB cores is the composite FR-4, which consists of epoxy resin, fiber glass reinforcements, and other chemical substances (based on bromine (Br)) to give the compound a flame retardant (FR) characteristic. It has a good liability for copper conductive path, a good mechanical stability, good electric creep resistance, and a low hygroscopicity. FR-4 has a maximum operating temperature suitable for industrial applications (115°C to 140°C), is adequate for up to high working frequencies, and has a relative permittivity $\varepsilon_r =$ 3.8 to 4.5. Core plates with thickness between 0.2 to 20 mm are usual.

A typical specification of the FR-4 double-laminated plate states strength of 1.5 mm and the copper layers having a 35-$\mu$m thickness.

Fewer disadvantages of the flat coil result in larger areas needing to get similar inductances. Due to the restricted available volume in a sensor, the first conclusion generally shows lower inductances for flat coils.

A practicable possibility to further increase the inductance is to use multilayer PCBs. They consist of more than two conductive layers (up to 48), which are isolated from each other. Identical coils can be printed on every layer and they can be connected in series by means of plated-through holes. This connection leads to a series inductance equal to the inductance of a single coil times the layer number (see (4.6)). In addition, the mutual inductances are higher at PCB coil. It can take advantage of the greater mutual inductances to increase the total inductance. In the experience of the author, the total inductance of the double-sided implementation is about four times the single inductance (compare with (4.9)). Of course, the effect is fortunately cumulative for more coil layers.

On the other hand, multilayer coils have higher losses, due to the holes.

*Example 5.4* This gives the reader the possibility to compare the parameters of the core coil in Example 4.4 (see Section 4.3.4 and Figure 3.20) with those of a PCB coil without a core having a comparable external diameter (9 mm).

The multilayer PCB inductor consists of 10 coils connected in series. Every single coil has the standard Archimedean spiral shape with 10 turnings. Its measured electrical parameters (see Figure 4.1) under undamped conditions (no target in front of the active face equivalent to $d = \infty$) are now:

$$R_{S\infty} = 1.4\Omega,\ L_{\infty} = 4.5\ \mu\text{H} \Rightarrow Q_{L\infty} = 50.5 \text{ and } R_{P\infty} = 790\Omega \text{ at } 2.5\text{ MHz} \tag{5.39}$$

The comparison with the values in (4.77) is distinct: the miniature PCB coil has lower losses ($R_{s\infty}$ is 10 times lower) and very small inductance ($L_{\infty}$ is 50 times lower). At any rate, the PCB-coil can work at higher frequencies and a comparable Q-factor can be achieved at 2.5 MHz.

The aforementioned fabrication procedure refers to single-sided (one copper layer) or double-sided PCBs (two copper layers on both sides of the substrate).

For multilayer structures, the manufacturing process is much more complex and correspondingly much more expensive. It is essentially a repetition of the fabrication steps listed above.

A resulting carrier core (after step 7) represents the inner layer structure and is stacked up double-sided with additional laminated copper layers isolated by prepregs (preimpregnated fibers). The next pressure-grouting step provides a multilayer lamination of the package. After intermediate steps, such as the drilling of reference holes by x-rays and plating through, similar steps as previously mentioned (exposure and development, copper layer processing, dry film stripping, etching, etc.) are repeated on the external layers.

For electrical connections between different layers, laser drilling microvias are used (thru-, blind-, step-, stacked-, buried-vias). There are many benefits to laser drilling microvias, including PCB area savings, impedance control, and circuit reliability. Diameters of microvias holes ($\geq 75\ \mu$m) are limited by aspect ratio (hole-depth to hole-diameter). As the hole-depth exceeds the hole-diameter, laser-drilling quality can be affected.

In opposition to thin laminates with cores above 50 $\mu$m, for industrial application under harsh climates, thick PCB circuits with thickness up to 20 mm and a thick copper layer can be fabricated.

Flexible structures (Flex PCB) on a polyimide base, which could have up to 10 layers, are also frequently used (e.g., electronic units manufactured by IFM electronic GmbH/Germany). They have a typical thickness of 0.17 mm and layer thicknesses of below 30 $\mu$m. One layer finally performs a shielding function. After the bonding of electronic surface-mount device (SMD) components on the foil, the sensor unit is coiled up with the shield layer as external sheet. A compact, functional electronic that has good EMC properties (see Section 2.2) results.

For sensor applications in HF ranges (very high frequency (VHF): 30 to 300 MHz, ultrahigh frequency (UHF): 300 to 3000 MHz, and super high frequency (SHF): 3 to 30 GHz), which require low losses of the carrier substrate material, the classical FR-4 can be replaced by better materials, such as (for example) laminates of the line RO4000® (manufactured by Rogers Corporation).

The popular product *Rogers RO4003C™* is a serial glass-reinforced hydrocarbon and ceramic (not polytetrafluoroethylene (PTFE)) laminates designed for performance sensitive, high volume commercial applications. It is designed to offer superior HF performance and low-cost circuit fabrication.

The established electrical parameters are a relative permittivity of $\varepsilon_r = 3.55$ at 10 GHz, and a loss factor $\tan\delta \leq 0.0027$. The semifinished panels have thicknesses between 0.2 to 0.5 mm and a double- or multilayer structure with a 35-$\mu$m copper thicknesses. A typical version has three copper layers in an alternating stack (copper–insulator–copper) of 35 $\mu$m–200 $\mu$m–35 $\mu$m–200 $\mu$m–35 $\mu$m.

In the author's opinion, this material is the best option to fabricate planar flat coils. In addition, Rogers technology offers the TICER™ TCR® option to integrate thin film resistor foil (NiCr, CrSiO, etc.) for resistances between 1Ω and 1 MΩ.

A nice illustration of miscellaneous PCB coils is shown in Figure 5.10. It is the evaluation panel LDCCOILEVM (about 145 mm × 140 mm), manufactured by Texas Instruments, which contains a large spectrum (of different shapes and sizes) of PCB test coil samples [7].

The majority of samples (12 pieces) is round coils (real Archimedean spirals with external diameter between 5 and 46 mm [see Figure 5.11(a)]) and can be used for target contactless actuation (see Section 1.2.1) by its axial approaching (see Section 2.1.1.1). The metallic target moves perpendicular to the coil surface maintaining its reference point on the coil rotational symmetry axis. The archived inductances cover the range 15 $\mu$H to 1 mH.

On the TI panel, there are also seven elongated rectangular PCB coils (length between 10 and 110 mm [see Figure 5.11(b)]), which can be used for target contactless actuation (see Section 1.2.1) by its lateral approaching for longitudinal travels (see Section 2.1.1.1). The metallic target provides a translatory motion along the

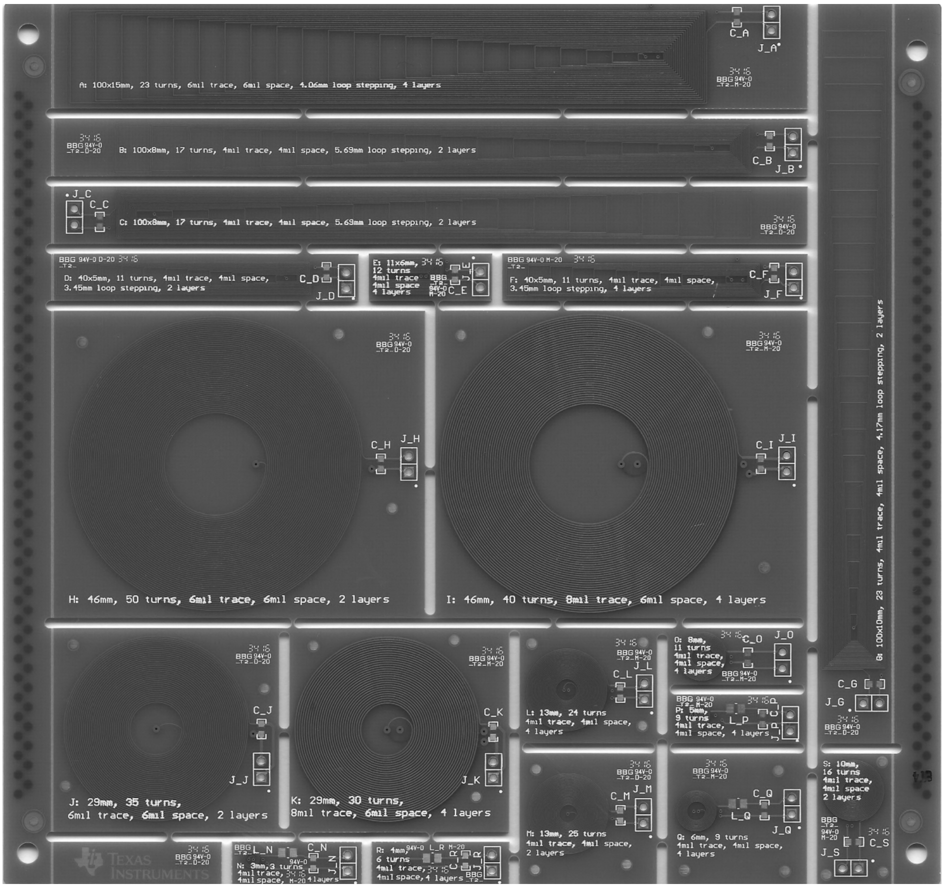

A: 100x15mm, 23 turns, 6mil trace, 6mil space, 4.06mm loop stepping, 4 layers
B: 100x8mm, 17 turns, 4mil trace, 4mil space, 5.69mm loop stepping, 2 layers
C: 100x8mm, 17 turns, 4mil trace, 4mil space, 5.69mm loop stepping, 2 layers
D: 40x5mm, 11 turns, 4mil trace, 4mil space, 3.45mm loop stepping, 2 layers
E: 11x6mm, 12 turns 4mil trace 4mil space 4 layers
F: 40x5mm, 11 turns, 4mil trace, 4mil space, 3.45mm loop stepping, 4 layers
G: 100x10mm, 23 turns, 4mil trace, 4mil space, 4.17mm loop stepping, 2 layers
H: 46mm, 50 turns, 6mil trace, 6mil space, 2 layers
I: 46mm, 40 turns, 8mil trace, 6mil space, 4 layers
J: 29mm, 35 turns, 6mil trace, 6mil space, 2 layers
K: 29mm, 30 turns, 8mil trace, 6mil space, 4 layers
L: 13mm, 24 turns 4mil trace, 4mil space, 4 layers
M: 13mm, 25 turns 4mil trace, 4mil space, 2 layers
N: 3mm, 3 turns 4mil trace, 4mil space, 4 layers
O: 8mm, 11 turns 4mil trace, 4mil space, 4 layers
P: 5mm, 9 turns 4mil trace, 4mil space, 4 layers
Q: 6mm, 9 turns 4mil trace, 4mil space, 4 layers
R: 4mm, 6 turns 4mil trace, 4mil space, 4 layers
S: 10mm, 16 turns 4mil trace, 4mil space 2 layers
C_A
J_A
C_B
J_B
C_C
J_C
C_D
J_D
C_E
C_F
J_F
C_G
J_G
C_H
J_H
C_I
J_I
C_J
J_J
C_K
J_K
C_L
J_L
C_M
J_M
L_N
C_N
C_O
J_O
L_P
C_Q
L_Q
J_Q
L_R
C_S
J_S
TEXAS INSTRUMENTS

**Figure 5.10** The evaluation panel LDCCOILEVM manufactured by Texas Instruments contains a large spectrum of PCB test coils [7].

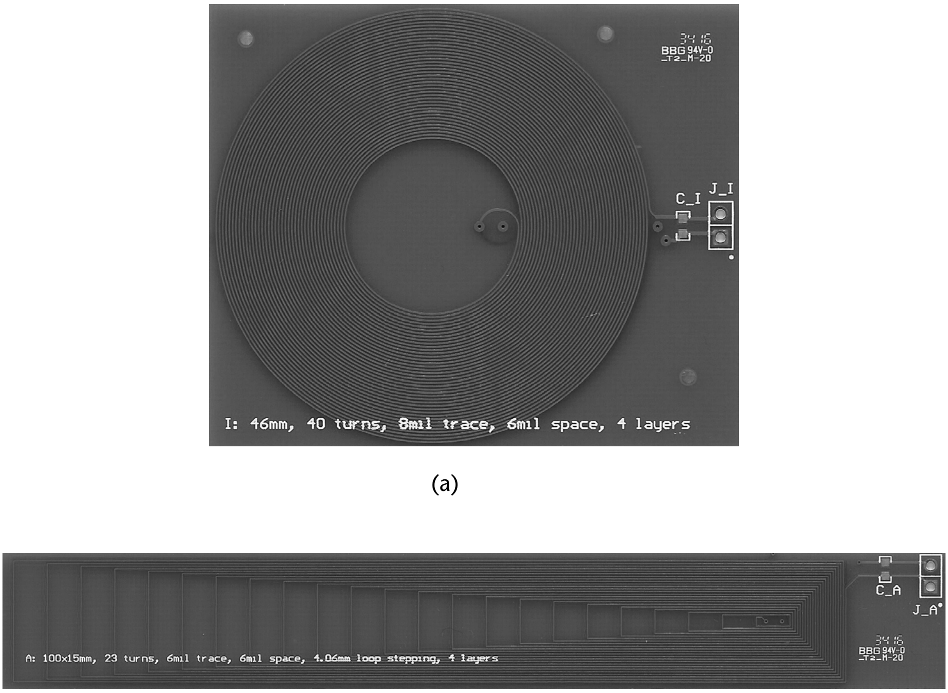
coil symmetry axis, keeping a constant distance between its surface and the coil surface. The realized inductances cover the range of 30 to 350 $\mu$H.

The flat PCB coils are generally used without any form of magnetic cores. However, it is necessary to have magnetic core, ferrite plates and—more preferred—permalloy, Mu-metals (see Chapter 6), or similar foils can be attached to the substrate.

Similarly to the calculation of the inductance of wired-wound coils (see Section 5.2), the internet provides many $L$ calculators for flat wired coils or planar spiral inductors (circular, octagonal, hexagonal, or square spirals).

Astonishingly, these tolls also generally go back to old paper [1], which still remains a classical reference in this field after a long time. Actually, a simple 90° rotation of the wired coil in Figure 5.3(c) metamorphoses this arrangement in a flat coil (see Figure 5.12). Equation (5.8) can be taken over and adapted to the new model. The result is a similar equation (see (5.40) in Table 5.3), although with adapted intermediate sizes. Equation (5.41) is another used calculation formula.

Some people tried to increase the computation accuracy going new ways. The reference paper [6] describes positive results achieved by means of three methods:


I: 46mm, 40 turns, 8mil trace, 6mil space, 4 layers
C_I
J_I
BBG 94V-0
A: 100x15mm, 23 turns, 6mil trace, 6mil space, 4.06mm loop stepping, 4 layers
C_A
J_A

(a)

(b)

**Figure 5.11** Example of (a) round, and (b) rectangular PCB coils from evaluation panel LDCCOILEVM, manufactured by Texas Instruments [7].

1. Modification of the classical expression developed by Wheeler [1];
2. Approximate computation derived from electromagnetic principles by approximating the sides of the spirals by symmetrical current-sheets of equivalent current densities;
3. Monomial expression derived from fitting to a large database of inductors.

The expressions were thoroughly verified, and they are accurate (typical errors of 2 to 3%), very simple, and suitable for use in design. The thickness of the inductor has only a small effect on inductance, and is therefore ignored.

The improved (5.42) uses the second-order geometric parameters of average coil diameter and fill ratio (see the legend in Table 5.3) and two correction coefficients $K_1$ and $K_2$, which slightly varies with the geometry. The formula fits full inductors ($D_o >> D_i$) but also hollow coils ($D_o \approx D_i$).

The approximate computation method in the list above leads to (5.43) in Table 5.3. A square spiral implies four identical current sheets. The current sheets in opposite sides are parallel to each other, whereas the adjacent ones are orthogonal. The

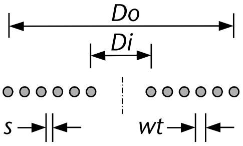


**Figure 5.12** Specification of a flat PCB coil ($D_o$: outer diameter, $D_i$: inner diameter, $s$: turn spacing, and $wt$: turn width).

**Table 5.3** Calculation table of the flat PCB coils

<table>
<tr><th><em>Internet Calculation Formula Derived from [1] (Require Units in Inches)</em></th><th><em>Computation Expressions According to the Methods Described in [6]</em></th></tr>
<tr><td>$$L = \frac{n^2 \cdot r^2}{8 \cdot r + 11 \cdot wc} \quad [\mu\text{H}] \qquad (5.40)$$ where: $wc = (s + wt) \cdot n$<br>and: $r = wc + Di/2$</td><td>$$L = \mu_0 \cdot K_1 \cdot \frac{n^2 \cdot Davg}{1 + K_2 \cdot \rho} \quad [\mu\text{H}] \qquad (5.42)$$ with:<table><tr><th><em>Layout</em></th><th>$K_1$</th><th>$K_2$</th></tr><tr><td>Square</td><td>2.34</td><td>2.75</td></tr><tr><td>Hexagonal</td><td>2.33</td><td>3.82</td></tr><tr><td>Octagonal</td><td>2.25</td><td>3.55</td></tr></table></td></tr>
<tr><td>$$L = \frac{n^2 \cdot A^2}{30 \cdot A - 11 \cdot Di} \quad [\mu\text{H}] \qquad (5.41)$$ where: $A = [Di + (s + wt) \cdot n]/2$</td><td>$$L = \frac{\mu_0 \cdot n^2 \cdot Davg \cdot C_1}{2}\left(\ln\frac{C_2}{\rho} + C_3\rho + C_4\rho^2\right) \qquad (5.43)$$ with:<table><tr><th><em>Layout</em></th><th>$C_1$</th><th>$C_2$</th><th>$C_3$</th><th>$C_4$</th></tr><tr><td>Square</td><td>1.27</td><td>2.07</td><td>0.18</td><td>0.13</td></tr><tr><td>Hexagonal</td><td>1.09</td><td>2.23</td><td>0.00</td><td>0.17</td></tr><tr><td>Octagonal</td><td>1.07</td><td>2.29</td><td>0.00</td><td>0.19</td></tr><tr><td>Circle</td><td>1.00</td><td>2.46</td><td>0.00</td><td>0.20</td></tr></table>$Davg = 0.5\ (Do + Di)$—average diameter<br>$\rho = (Do - Di)/(Do + Di)$—fill ratio</td></tr>
</table>

$n$ = number of turns; $s$ = turn spacing; $wt$ = turn width; $r$ = coil radius; $wc$ = coil width.

authors use symmetry properties and the fact that parts with orthogonal current sheets have zero mutual inductance. The inductance computation is now reduced to evaluating the self-inductance of one sheet and the mutual inductance between opposite current sheets. For this evaluation, the authors use the concepts of geometric mean distance, arithmetic mean distance, and mean square distance. The resulting expression contains four geometry coefficients, $C_1$ to $C_4$, and provides a good accuracy for coils with small spacing ($s < 3 \cdot wt$). This condition is fulfilled by practical PCB spiral coils, which generally have a spacing $s \leq wt$.

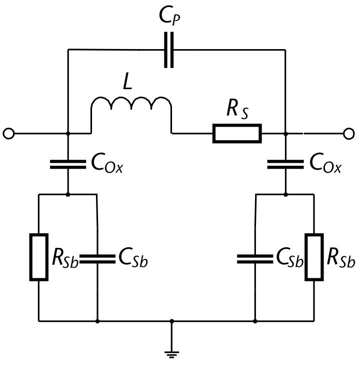


**Figure 5.13** Equivalent circuit of a flat PCB coil at HFs.

For working at HFs, the equivalent circuit in Figure 4.1 is only of limited suitability. The schematic has to be extended by the coil parallel capacitance $C_P$ and the shunt path, which consists of the adhesive (oxide) capacitance $C_{Ox}$ and of the substrate parasitic components $R_{Sb}$ in parallel to $C_{Sb}$ (see Figure 5.13) [8]. At these frequencies, the capacitances $C_{Ox}$ are effective and cross currents flow through the lossy shunt paths. The influence can be reduced using good coil substrates (e.g., Rogers RO4000® materials). Furthermore, the capacitance $C_P$ between spiral turns is also remarkable.

The preponderant use of flat coils, already printed on PCB and immediately connected to vicinal bonded SMD electronic components, to provide a reactive part refer to local frequency generators as pilot oscillators, clock generators, and so forth.

Making abstraction of such traditional applications, the use of stand-alone PCB coils in ISE is concentrated on two major application areas:

1. *Sensing coil* for IPS (see Section 2.1), which perform a contactless actuation (see Section 1.2.1) of a metallic target by evaluation of the eddy currents (see Section 3.3.6). The traditional embodiment for this purpose is the wire-wound coil in a ferrite pot core (see Figure 3.20).

   At any rate, an attentive study of the novel, high performance products in this market discovers the implementation of the flat PCB instead of the conventional coil. An informative paper [9] explains how the global player, namely Hans Turck Company, develops and manufactures the so-called Factor 1 IPSWs (see Section 2.1). In contrast to standard IPSW, these have rated operating distances ($s_n$) (see Figure 2.3) that are independent of the target metal. For this purpose, the sensing element consists of three PCB coils placed parallel to each other in the frontal range of the sensor. The sending coil is placed in the middle of two receiving coils. There is a differential method of measurement. The sending coil generates eddy currents in the target, and these induce voltages in receiving coils. Due to their different position relative to the target, the back-induced signals in these are different in value and phase. The sensor electronics evaluate these differences and give the expected switching behavior. This piggyback board arrangement is possible only due to the PCB coil technology.

   The patent-protected procedure was further developed to achieve Factor 1, but also high operating distances. The sensor contains two pairs of both sending and receiving coils, respectively. Apart from that, the evaluation remains similar.

   Certain companies prefer simply to maintain the standard evaluation procedure of the IPSW than to replace the wired coil with a PCB coil. In order to improve the sensor stability, the coil terminal voltage is compared with the terminal voltage of a second dummy coil connected in series, but placed in the sensor inside and not influenced by the target. A key to success is to have two coils as identical as possible. The patent of the global operating player, namely Balluff Company, stated to initially manufacture the sensor PCB with two coils placed side by side hereon. Before the bonding of electronic parts starts, one coil is cut out from sensor PCB and stub connected (mechanically and electrically) by perpendicularly soldering on sensor PCB. The second coil remains integrated in the PCB and so has a perpendicular position to the first one.

2. The application area includes novel ISE, of modern IS types, where the PCB topology is a prerequisite condition, for example, at Sagentia™-based sensors.

## 5.5 Integrated Coils on Silicon Substrate

The development and fabrication of the coils on silicon, also named coil-on-chip, began in the nineties as a progression of the significative, racy growth of the complementary metal-oxide-semiconductor (CMOS) technologies for the manufacturing of integrated circuits (ICs) and ASICs (see Section 7.3). This space- and money-saving technique provides monolithic integration of CMOS transistors: both p-channel and n-channel on the common silicon substrate, but also structuring of passive R-C-L components. If the integration of R and C was a takeover from the previous bipolar technology of the sixties through the eighties, the coil structuring on silicon substrate represented a new challenge. The challenge primarily refers to the miniaturized sizes (below a few millimeters) and the needed accuracy (see Figure 5.14). The implemented geometries and fabrication represented an improved transfer from the PCB coil topologies and photolithographic methods (see Section 5.4).

The need of small and reliable coils placed much closer to the electronics (on the same chip or aside) was pushed by large scale applications such as communications, bank cards, radio frequency identification (RFID) systems, telemetry (tag antenna, local oscillator, etc.).

Development projects of coils-on-chip for implementation as ISEs were much rarer and performed by research institutes (some in Switzerland) in cooperation with interested companies [10–12].

Involving the case study method, the recent volume [13] explains what the technical challenges to form a useful inductor on a silicon substrate with only a thin metallization layer available are, such as in standard planar silicon process

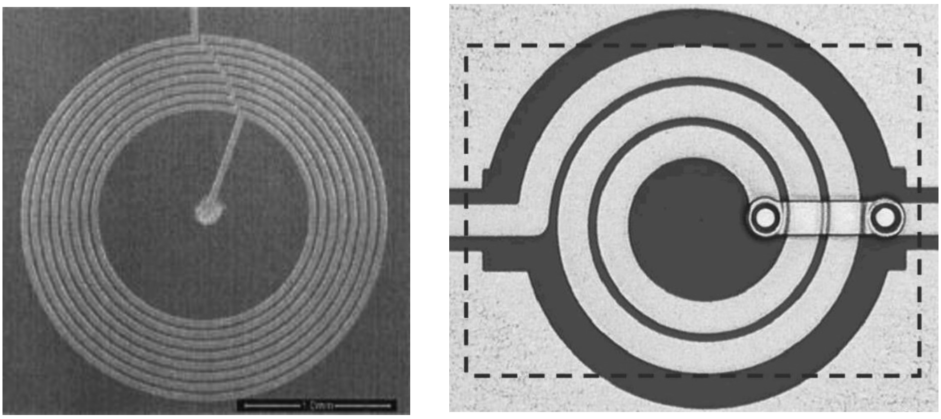

(a) (b)

Figure 5.14 (a) Scanning electron microscope picture of planar microcoil fabricated in silicon using a simple photolithography (radius 750 $\mu$m, single metal layer), and (b) planar inductor integrated on p+ type porous silicon regions and laterally shielded by the metal layer (diameter 500 $\mu$m).

technologies. The authors manage with a similar equivalent circuit as in Figure 5.13 (yet $C_{Sb}$ is removed) and underline the following specific features and concerns:

- A particular concern is that coils for high inductances are relatively large and lead to higher costs and have lower values of the Q-factor. Unfortunately, they are demands for ISE:
- The linear frequency sweep of the coil Q-factor $Q_L$, which is dependent on the main coil parameters $L$ and $R_S$ (see (4.27)) is valid up to a certain frequency $f_{\max}$. Above this point, the shunt path (parasitic capacitance $C_{Ox}$ between metal and substrate in series with the bulk resistance $R_{Sb}$) overlays the Q-factor, increasing with a falling behavior:

$$Q_C \approx \frac{1}{2\pi \cdot f \cdot C_{OX} R_{Sb}} \tag{5.44}$$

  The frequency $f_{\max}$ has to be adjusted higher than the chosen working frequency range by tailoring the terms $2\pi fL/R_S$ and $1/2\pi fC_{Ox} \cdot R_{Sb}$;
- The parallel capacitance $C_P$ causes an undesirable parallel self-resonance, whose resonance frequency $f_{SR}$ should be set considerably higher by optimizing of the interwire capacitance of the layout:

$$f_{SR} = \frac{1}{2\pi\sqrt{LC_P}} \tag{5.45}$$

- The losses resistor $R_S$, which lowers the Q-factor, depends on the resistive metallization component and also on the magnetic field that extends into the substrate and causes eddy currents in the conductive silicon.

  The losses resistor $R_S$ can be minimized by using CMOS-processes with multi-metal layer features and connecting these layers in parallel;
- The substrate resistivity plays an important role in the tuning of the coil features. For the substrate resistivity above of about 0.1 Ωm, eddy currents and capacitive losses decrease as resistivity increases, resulting in higher $Q_L$ values (inductor mode). For the substrate resistivity below of about 0.01 Ωm, eddy currents rise so high in substrate that $L$ becomes reduced (eddy current mode). For values between these limits, the frequency visibly drops and the coil acts as a resonator (resonator mode) with:

$$Q = \frac{Q_L \cdot Q_C}{Q_L + Q_C} \tag{5.46}$$

  To prevent unwanted modes, a uniform metal shield should be interleaved between the coil structure and the substrate. Possible eddy current losses in this shield can be controlled by patterning measures.
- In summary, [13] reports that a competitive coil with a Q-factor of 20 was achieved using copper metallization instead of the classical aluminum metallization and high-resistivity silicon substrate. For much higher Q-factor values and minimum area, solenoid-like stacked coil structures are recommended.

The Institute of Microsystems and Microelectronics at the Swiss Federal Institutes of Technology (EPFL), Lausanne is one of the representative participants, together with known electronic manufacturers in Switzerland (Baumer Electric AG, Asulab Division of Swatch Group SA), focused on the efforts to make coils-on-silicon available for ISs. A successful implementation is reported in [11]. It concerns a prototype of a highly miniaturized (3.5 mm × 3.5 mm × 1.2 mm), fully integrated inductive proximity microsensor. The sensor is composed of a planar sensing coil on a substrate and evaluation electronics.

Not only are the sensor electronics fully integrated, but also the coil of the sensing element (Figure 1.2). Low-fabrication costs (standard 0.8 $\mu$m CMOS process), robustness, and a good stability over temperature (despite silicon coil) were focused on during the development.

To evaluate the influence of different geometrical parameters, working conditions, and configuration and materials, a test integration of a dummy testing wafer was initially started. A variety of flat spiral coils (circular and square) having diameters between 100 $\mu$m and 10 mm, with copper turns of 25-$\mu$m thickness and width between 30 and 60 $\mu$m have been electroplated. After the experience accumulated by the benchmarking of these coils, a coil geometry for the above mentioned sensor was decided, namely a suitable electroplated sensing copper coil with the side length of 3.5 mm, pitch of 25 $\mu$m (20/5) and copper thickness of 25 $\mu$m. The typical coil parameters with any target influence resulted in $L = 4.3\ \mu$H and $R_S = 18\Omega$ (see Figure 4.1) at a working frequency of 25 MHz. Despite the relative high value of $R_S$ (thin copper structure), the Q-factor achieves practicable values ($Q_L = 37.5$ at 25 MHz) and confers applicability for inductive proximity jobs.

The sensor electronics are based on a relaxation oscillator with an op-amp in differential configuration (see Figure 8.41(b)), which generates a squared AC output signal, whose frequency depends on the proximity of the target to the sensing coil. The approach of the conductive target changes the magnetic field distribution, increases the losses, modifies the coil time constant $\tau = L/R_S$ (see (4.23)), and finally drops the output frequency (down to about 10 MHz.) For an aluminum target (1-mm thickness), the sensor exhibits an SR of 1500 $\mu$m with a linearity range $LR$ = 100 to 700 $\mu$m (see Figure 1.3) and a resolution above of 0.1 $\mu$m (see Figure 1.6). The typical frequencies of the sensor output, free from target and close to the target, are 30 and 10 MHz respectively. A significant tuning point of the output characteristic is 24 MHz at the distance $s_r = 500\ \mu$m (see Figure 2.3).

Due to an ingenious temperature compensation of the temperature drift of the sensing coil, the sensor works with small temperature deviation of the output signal in the wide industrial temperature range of −40°C to +80°C.

Specific for the demands of industrial applications, the sensor is suitable for unipolar supply voltage (0–5V). Three-wire supply topologies ($+V_{DD}$, GND, $-V_{SS}$) are not acceptable in most industrial applications or very difficult to adapt (see Chapter 11). The low power consumption is below 100 mW in operation mode and below 1 mW in power-down operation, which disables the electronic part.

The sensor packaging has a compact sandwich configuration. The electronic chip (die-form without package) is attached faceup on the upper side of a 100-$\mu$m thin PCB substrate. On the other side of this PCB is the flat coil placed facedown. Its terminal connections are conducted by vias from the active coil face to the bonding pads placed on the back side. The connections between chip input pads and these coil back side pads, as well as the connections between the chip output and supply pads and sensor application pads on the carrier PCB, are made by a wire-bonding procedure (see Section 7.3.1.2). Finally, the chip area is covered by a coating protective material (glob top encapsulation). The active face of the sensing coil remains free from any obstacles, which allows a nearer access of the target to the coil.

The hybrid configuration of the microsensor composed of a sensor electronics chip with small dimensions (below 1 mm × 1mm) and of a sensing coil on separated substrate with a surface 10 times greater that the electronics surface was preferred from the point of view of sensor costs (silicon saving).

Finalizing this section, it can be concluded that contrary to the topic coil on PCB, which is well-established in the field of ISs, the subject matter of coil-on-chip has, at the moment, an academic significance. Implementations in serial production seem to be unknown, but their time will come.

## 5.6 Active Inductors, Gyrators

Gyrators are not real parts of the ISE family and are therefore not used. Nevertheless, these electronic circuits are a good and flexible replacement for real inductors in test jigs or in automated test systems for electronic units of an IS. That is the reason why the gyrator is included in this section.

Active inductors are an application of the conventional active gyrator circuitry and can be used as equivalent active circuitry to a passive element (capacitor or inductor).

A gyrator is an active, linear, lossless, two-port electrical network element proposed in 1948 by Bernard D. H. Tellegen (who also proposed the circuit symbol, see Figure 5.15), which can transform an arbitrary impedance in its dual impedance. A gyrator inverts the current-voltage and impedance characteristics of linear electrical elements. Thus, a gyrator can make a capacitive circuit behave inductively, a series LC circuit behave like a parallel LC circuit, and so on.

Important for our book, the gyrator can provide an inductive input behavior if a capacitive load is connected to its output.

The transfer equations of the ideal gyrator are [14]:

$$\underline{I}_1 = 0 \cdot \underline{V}_1 + \frac{1}{R_g} \cdot \underline{V}_2 \tag{5.47}$$

$$\underline{I}_2 = \frac{1}{R_g} \cdot \underline{V}_1 + 0 \cdot \underline{V}_2 \tag{5.48}$$

where $R_g$ is the gyration resistance of the gyrator.

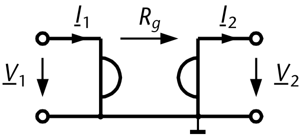


**Figure 5.15** Gyrator schematic labeled.

If a load impedance $\underline{Z}_L$ is connected at the output, the input impedance $\underline{Z}_{\text{inp}}$ results according to the two-port network theory:

$$\underline{Z}_{\text{inp}} = \frac{R_g^2}{\underline{Z}_L} \tag{5.49}$$

If the load particularly is a capacitor $C$, (5.49) changes showing the above-mentioned $C$ to $L$ transformation:

$$\underline{Z}_{\text{inp}} = \frac{R_g^2}{\dfrac{1}{j\omega \cdot C}} = j\omega \cdot C \cdot R_g^2 \tag{5.50}$$

These proportionalities between a current of one side (input/output) and the voltage of the other side (output/input) can faithfully be reproduced through circuitry using voltage-controlled current sources with high input and output impedances, respectively (Figure 5.16).

Tellegen suggested a number of ways in which a practical gyrator might be built. At low frequencies and low powers, circuits that operate as gyrators can be built with transistors and—timely—with small op-amps. Reference [14] suggests using the transconductance amplifier OPA615 (Texas Instruments), which works in a broad frequency band and provides very high output impedance (MΩ-range) due to its cascade-current source at the output.

Figure 5.17 shows classical active gyrator circuits with op-amps and their (approximate) equivalent passive circuits [15].

In the circuit shown in Figure 5.17(a), one port of the gyrator is between the input terminal and ground, while the other port is terminated with the capacitor. The circuit works by inverting and multiplying the effect of the capacitor in an $RC$ differentiating circuit where the voltage across the resistor $R$ behaves through time

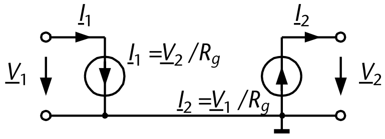
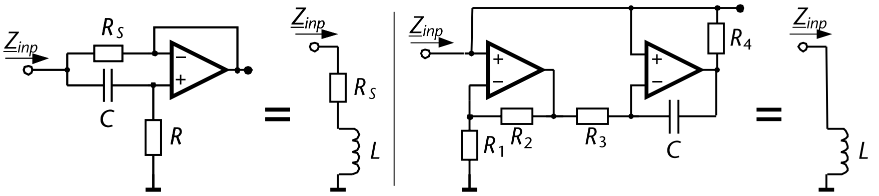

$\underline{I}_1$
$\underline{I}_2$
$\underline{I}_1 = \underline{V}_2 / R_g$
$\underline{V}_1$
$\underline{V}_2$
$\underline{I}_2 = \underline{V}_1 / R_g$

**Figure 5.16** Implementation of a gyrator using two voltage-controlled current sources.



(a) (b)

**Figure 5.17** Replication of (a) a real inductor, and (b) an ideal inductor.

in the same manner as the voltage across an inductor. The op-amp follower buffers this voltage and applies it back to the input through the resistor $R_S$.

Thus, the input impedance $\underline{Z}_{\text{inp}}$ is given by a parallel connection of two paths. One of these is the series path $R$ and $C$. In typical designs, $R$ is chosen to be sufficiently large ($R >> R_S$) so that the other term dominates. Thus, the RC circuit's effect on the input impedance is negligible. The expression of input impedance $\underline{Z}_{\text{inp}}$ denotes an impedance similar to (4.12) with an adjustable real part $R_S$ and an inductive part $L$:

$$\underline{Z}_{\text{inp}} = R_S + j\omega L = R_S + j\omega \cdot R_S \cdot R \cdot C \tag{5.51}$$

This circuitry is the typical application of the high precision op-amp MCP6051 (Microchip Technology, Inc.). There is a practical limit on the minimum value that $R_S$ can take, determined by the current output capability of the op-amp. Recommended range is $R_S = 1$ to $100\Omega$.

The circuitry in Figure 5.17(b) theoretically provides an impedance:

$$\underline{Z}_{\text{inp}} \approx j\omega L = j\omega \cdot \frac{R_1 \cdot R_3 \cdot R_4 \cdot C}{R_2} \tag{5.52}$$

which approximates an ideal inductor. Usually, the four resistors have the same value.

Simulated elements are electronic circuits that imitate actual elements. They cannot replace physical inductors in all the possible applications, as they do not possess all the unique properties of physical inductors. Resulting features of the coil simulation by gyrators are:

- Positive attributes of the gyrator include that it has no moving parts and can be implemented in a fairly compact PCB.
- Gyrators can be used to create inductors in a wide range, from microhenry up to megahenry values. Physical inductors are typically limited to tens of henries, and have parasitic series resistances from hundreds of microoohms through the low kiloohm range. The series resistance of a gyrator depends on the topology and typically ranges from ohms to hundreds of ohms.
- Physical capacitors today are much closer to ideal capacitors than physical inductors are to ideal inductors. Because of this, a synthesized inductor realized with a gyrator and a capacitor may be closer to an ideal inductor than any physical inductor can be. The Q-factor of a synthesized inductor results higher and can easily be adjusted.
- Gyrator inductors typically have higher accuracy than physical inductors, due to the lower cost of precision capacitors over inductors.
- Gyrator inductors can carry on doing resonance series or parallel circuits (see Figure 4.6). An example of such implementation and of the equivalent circuit is shown in Figure 5.18.
- Gyrator inductors are simulated inductors and are not real inductors. They do not have the inherent energy storing properties of the real inductors and this limits the possible power applications. The circuit cannot respond like a real inductor to sudden input changes (it does not produce a high-voltage back

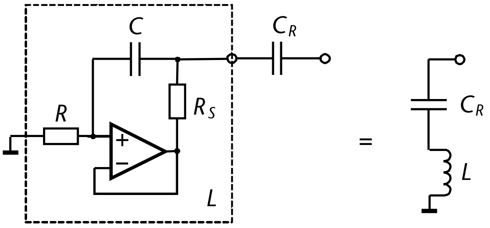


**Figure 5.18** Series resonant circuit reproduced by means of a gyrator inductor.

emf). They are not useful for modeling of the flyback property of inductors, where a large voltage spike is caused when the current is interrupted.

- Since gyrators use active circuits, they only operate as a gyrator within the power supply range of the active element. The voltage response is limited by the power supply. A gyrator's transient response is limited by the bandwidth of the active device in the circuit and by the power supply.
- The fact that one side of the simulated inductor is grounded restricts the possible applications (real inductors are floating). However, the gyrator can be used in a floating configuration with another gyrator so long as the floating grounds are tied together. This allows for a floating gyrator, but the inductance simulated across the input terminals of the gyrator pair must be cut in half for each gyrator to ensure that the desired inductance is met (the impedance of inductors in series adds together).
- Simulated inductors do not react to external magnetic fields, and permeable materials the same way that real inductors do. In addition, they do not create magnetic fields and induce currents in external conductors the same way that real inductors do.

In summary, these attributes limit the use of gyrator inductors in applications such as sensors, detectors, and transducers.

## References

[1] Wheeler, H. A., *Simple Inductance Formulas for Radio Coils*, Boca Raton, IEEE Publisher, 1928.

[2] Dorf, R. C. (ed.), *The Electrical Engineering Handbook*, Boca Raton, FL: CRC Press / IEEE Press, 1997.

[3] American Radio Relay League, *The ARRL Handbook for Radio Communication*, 83rd edition, Newington, CT: ARRL Publisher, 2006.

[4] Neosid Pemetzrieder GmbH & Co. KG, *Neoside_catalogue_2015_complete*.pdf: http://neosid.de, 2015.

[5] Asch, G., *Les Capteurs en Instrumentation Industrielle*, 7th edition, Paris, France: Dunod, 2010.

[6] Mohan, S. S., et al., "Simple Accurate Expressions for Planar Spiral Inductances," *IEEE Journal of Solid-State Circuits*, Vol. 34, No. 10, October, 1999.

[7] Texas Instruments, *Application Report SNOA954: LDC Device Selection Guide*, July 2016.

[8] Muehlhaus, V., "Entwurf und Simulation von RFIC Spulen," *HF-Praxis*, Germany, March, 2015.

[9] Hans Turck GmbH & Co. KG, *Grundlegende Informationen zu induktiven Sensoren*, MessTec & Automation, December, 2005.

[10] Sadler, D. J., C. H. Ahn, "On-Chip Eddy Current Sensor for Proximity Sensing and Crack Detection," *Sensors and Actuators A: Physical*, Elsevier, Vol. 91, No. 3, 2001.

[11] Kejik, P., et al., "A Low-Cost Inductive Proximity Sensor for Industrial Applications," *Sensors and Actuators A: Physical*, Vol. 110, No. 1–3, 2004.

[12] Wu, J., G. H. Bernstein, "A Microfabricated Transduction Coil for Inductive Deep Brain Stimulation," *IFSA Sensors & Transducers*, Vol. 69, No. 7, July, 2006.

[13] J. N. Burghartz (ed.), *Guide to State-of-the-Art Electron Devices*, John Wiley & Sons, 2013.

[14] Tietze, U., C. Schenk, and E. Gamm, *Halbleiter-Schaltungstechnik*, 15th edition, Berlin Heidelberg, Germany: Spinger/Vieweg-Verlag, 2016.

[15] Hooper, R., B. Guy, and R. Perrault, "A Current-Controlled Variable Inductor," *IEEE Instrumentation & Measurement Magazine*, August 2011.

CHAPTER 6

# Magnetic Materials for Cores and Plungers of ISEs

Chapter 5 fully described how the inductors of the ISEs are made (round wired, printed, or integrated coils), how they can be designed and calculated, and which are the most often used materials and technologies for these coils.

The main classification of the inductors for ISEs distinguishes between coils without a core and coils with a magnetic core, which are fixed or movable (plunger). The benefits of the magnetic core were also defined and examined in the previous chapter. Finally, suitable magnetic materials for cores of ISEs were introductorily listed in Chapter 5.

Chapter 6 carries forward with the presentation of these magnetic materials, focusing on issues such as the definition of physical parameters and features, component elements and fabrication technologies (metallurgy), shapes and types (standard or application specific), among others.

## 6.1 Ferrites

Ferrites are nonmetallic, ceramic-like materials with magnetic properties [1], which are used in many types of electronic devices. A sintered ferrite consists of small crystals, typically 10 to 20 $\mu$m in dimension. The so-named Weiss domains exist within these crystals in which the molecular magnets are already aligned (ferrimagnetism). When a driving magnet field is applied to the material, the domains progressively align with it.

Ferrites are used in permanent magnets, cores for transformers and toroidal inductors, computer memory elements, solid-state devices, and last but not least, ferrite cores for ISs.

Ferrites are composed of ferric oxide having low electrical conductivity and one or more other metals in chemical combination. A ferrite is usually described by the formula $M\_Fe_x\_O_y$, where M represents any metal that forms divalent bonds, such as manganese (Mn), zinc (Zn), nickel (Ni), cobalt (Co), copper (Cu), or magnesium (Mg).

The material properties of ferrites include:

- Hard and brittle;
- Containing iron;
- Polycrystalline;
- Generally gray or black.

Physical properties of ferrites include:

- Significant saturation magnetization $M$ ($\approx B$, if $\mu_o H << B$ [see (4.50)]);
- Generally high electrical resistivity, lower conductivity;
- Low electrical losses;
- Very good chemical stability (chemically inert).

### 6.1.1 Ferrites: Classification, Definitions, and Properties

Ferrites are often classified as soft or hard in terms of their magnetic properties:

- *Soft ferrites* [2] are used in electromagnetic cores or transformers. The most popular combinations are manganese-zinc (MnZn) or nickel-zinc (NiZn) ferrites. These compounds exhibit a low coercivity, they can easily be magnetized and used up to VHFs.

  MnZn ferrites are predominantly used in ISE, can work at high frequencies up to the order of a few megahertzs, have high magnetic permeability and saturation induction, as well as low loses.
- *Hard ferrites* have a high coercivity. They are strontium, barium ferrites, and are widely used in household products.

Soft ferrites do not retain significant magnetization, whereas the magnetization of hard ferrites is considered permanent.

Ferrite components are pressed from a powdered precursor and then sintered (fired) in a kiln. The mechanical and electromagnetic properties of the ferrite are heavily affected by the sintering process which is time-temperature-atmosphere dependent. Ferrites significantly shrink when sintered. Depending on the specific ferrite, this shrinkage can range from 10 to 17% in each dimension. Maintaining correct dimensional tolerances as well as the prevention of cracking and warpage related to this shrinkage are fundamental concerns of the manufacturing process.

The ferrites used in the cores of the ISE of an IS are soft ferrites, which are characterized by following features (see Figure 6.1) [3]:

- High magnetic permeability ($\mu < 20{,}000$);
- Slim hysteresis loop ($B_r < 0.5$T) ;
- Low coercive field strength ($H_c < 100$ A/m);
- High saturation flux density ($B_s < 2$T).

Reference [1] is appreciated to be an invaluable resource for the IS research scientists and technicians, as well as users. The worldwide ranking of ferrite manufacturers and suppliers is comprised of more than ten well-known names. Four of these names [4–7] can be found at the end of this chapter. Very high practical and theoretical levels of knowledge are specific for their ferrite data sheets. An important part of this information is processed in this chapter.

After this introductory specification of the soft ferrites, a detailed presentation of the main ferrite parameters should be useful. A selection of the ferrite attributes (more than 30 overall), which are important for analyzing and the design of ISEs, leads to the following enumeration:

*1. Hysteresis loop (BH curve):* During the magnetization process, energy barriers have to be overcome. Therefore, the magnetization will always lag behind the field. A so-called hysteresis loop is the result of this behavior (see Figure 6.1). If the material has a small resistance against magnetization, a large induced flux will result at a given magnetic field, which is called high permeability. The BH curve is a graphical representation of the relationship between magnetic field strength $H = |\mathbf{H}|$ and flux density $B = |\mathbf{B}|$, defined as magnitude of the corresponding vectors **H** and **B** (see (4.46)).

Assume an alternating field is applied to a soft magnetic material after a preliminary complete demagnetization. Proceeding from $H = 0$ and $B = 0$ and increasing the magnetic field strength, the so-called initial magnetization curve (virgin curve) is first obtained. At the beginning, this curve consists of a straight line with a constant slope.

For very high $H$ values, the curve bends, taking lower slopes until the peak of maximum attainable flux density is reached. This point is known as the saturation flux density $+B_s$. It flags the limit of the BH curve, where further increasing of $H$ does not have any effect, and the $B$ value remains unchanged. The majority of ferrite suppliers technically specify the $B_s$ at a field strength $H = 1{,}200$ A/m.

If the field is removed or reduced, the working point evolutes now on the real BH curve. If the magnetic field reaches the zero value, a certain flux density remains. That is the material remanent flux density $+B_r$ (intersection with the ordinate $H = 0$).

The remanent flux returns to zero for a certain negative field strength. The intersection point with the abscissa $B = 0$ defines the negative coercive field $-H_c$.

By continuing to increase the negative field, the operation point ends up on the negative saturation flux density $-B_s$.

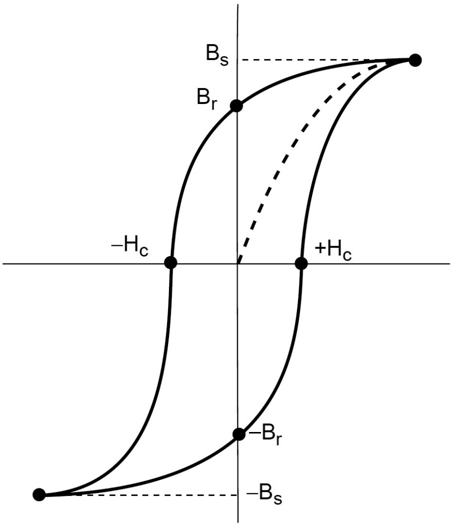


**Figure 6.1** The initial magnetization curve (dotted representation) and the associated hysteresis loop (solid curve): $B(\mu, H) = \mu_0 \cdot \mu_r(H) \cdot H$.

If the magnetic field now increases from the lowest negative value to the highest positive value, the operating point executes a similar movement on the other side of the BH curve passing the distinctive point: $-B_r$ and $+H_c$ and tending to the $+B_s$ peak.

2. *Initial permeability ($\mu_i$):* In previous chapters, the medium absolute permeability $\mu$ is defined from an electrodynamic point of view, namely as ratio of flux density and applied field (compare with (4.47)). This permittivity is expressed as the product of the magnetic constant of free space $\mu_0$ and the relative permittivity $\mu_r$ (e.g., (3.3)).

Since there are several versions of $\mu_r$ depending on conditions, the index $r$ is generally removed and replaced by the applicable symbol (e.g., $\mu_i$, $\mu_e$, and so forth). The initial permeability is measured in a closed magnetic circuit (ring core like in the lower left corner in Table 6.3) using a low magnetic field (B ≤ 0.25 mT) that corresponds to an operation point close to the origin of the BH curve:

$$\mu_i = \frac{1}{\mu_0} \cdot \lim_{H \to 0} \frac{\Delta B}{\Delta H} \tag{6.1}$$

Initial permeability is dependent on temperature and frequency. Ferrite providers specify a cutoff frequency $f_C$ as one where $\mu_i$ is halved from its value at low frequencies: 1 to 10 kHz (e.g., the F08 material has an $f_C = 5$ MHz [5]). According to IEC 60401-3, $\mu_i$ is defined using a toroid etalon at $f = 10$ kHz, B = 0.25 mT, and 25°C.

3. *Effective permeability ($\mu_e$):* Most core shapes in use today do not have closed magnetic paths or a constant geometry. If an air gap is introduced in a closed magnetic circuit, magnetic polarization becomes more difficult. As a result, the flux density for a given magnetic field strength is lower. The effective permeability is generally defined as the permeability of a hypothetical material, supposed to be homogenous, to get the same total remanence as for the real system consisting of parts with identical dimensions but with different permeability.

Making an air gap in a simple closed magnetic circuit, which has a constant cross-sectional area, the permeability $\mu_i$, and the effective length $l_m$ (compare with Figure 3.7), the resulting effective permeability $\mu_e$, can be approximated by:

$$\frac{1}{\mu_e} \approx \frac{1}{\mu_i} + \frac{l_0}{l_m} \tag{6.2}$$

where the length $l_0$ of the air gap is substantially shorter than $l_m$ (<0.5%) [4].

The introduction of an air gap in the magnetic path causes a clockwise rotation of the BH curve (permeability decreasing). For longer air gaps, some flux will cross the gap outside its normal area causing an important change of the effective permeability.

4. *Apparent permeability ($\mu_{app}$):* This characterizes the influence of the winding core if this occupies a limited area of the entire magnetic field, as in the case of the systems with open magnetic loop (see Section 3.3.4). The feature is defined as a ratio between the inductance $L_0$ of the coil without core and the inductance $L_C$ of the same coil with a core in a determined position:

$$\mu_{\text{app}} = \frac{L_C}{L_0} \tag{6.3}$$

The definition of $\mu_{\text{app}}$ is particularly important for coils with tubular, cylindrical, and threaded coils, since an unambiguous relationship between $\mu_i$ and $\mu_e$ is not possible because of the high leakage inductances. An accurate specification of the $\mu_{\text{app}}$ requires a precise specification of the measuring coil arrangement. A typical value for half pot cores implemented in IPS is $\mu_{\text{app}} = 5$ to 10.

5. *Incremental and reversible permeability ($\mu_\Delta$, $\mu_{rev}$):* In order to measure the incremental permeability $\mu_\Delta$, a relatively small measuring alternating field is superimposed on a static DC polarization field $H_{DC}$, which generally has a strong influence on permeability:

$$\mu_\Delta = \frac{1}{\mu_0} \cdot \left[\frac{\Delta B}{\Delta H}\right] \text{ at } H_{\text{DC}} \tag{6.4}$$

If the amplitude of the alternating field is negligibly small relative to $H_{DC}$, (e.g., ISE applications), the permeability is then called reversible permeability $\mu_{\text{rev}}$.

6. *Complex permeability ($\underline{\mu}$):* A real inductor will never be an ideal inductance but it has an equivalent series circuit (see Figure 4.1). For the impedance of this series representation (compare with (4.12)):

$$\underline{Z} = R_S + j\omega L \tag{6.5}$$

The magnetic losses are accounted for if a resistive term is added to the permeability. The complex permeability expressed by:

$$\underline{\mu} = \mu' - j\mu'' \tag{6.6}$$

where the relative real part corresponds to the initial permeability (inductance):

$$\mu' = \Re\{\underline{\mu}\} \tag{6.7}$$

and the relative imaginary part, called imaginary permeability:

$$\mu'' = \Im\{\underline{\mu}\} \tag{6.8}$$

characterizes the magnetic losses.

The introduction of complex permeability offers the possibility to express the impedance $\underline{Z}$ in a compact form [6]:

$$\underline{Z} = j\omega\underline{\mu} \cdot L_0 \tag{6.9}$$

where $L_0$ represents the inductance of a core of permeability $\mu_r = 1$, but with unchanged flux distribution.

Making the substitution (6.6) into (6.9), the factorized expression explains the significance of the complex permeability parts:

$$\underline{Z} = \omega\mu'' \cdot L_0 + j\omega\mu' \cdot L_0 \tag{6.10}$$

The frequency dependence of these two parts is sometimes offered by the ferrite producer [5] and permits determination of the magnetic loses at the working frequency. In Figure 6.2 the real component $\mu'$ of different ferrites is constant at low frequencies, attains a maximum at higher frequencies, and then drops in approximately inverse proportions to frequency. At the same time, the imaginary part $\mu''$ rises steeply from a very small value at low frequencies to attain a distinct maximum and, past this, also drops as the frequency is further increased.

*7. Loss factor:* The loss factor tan $\delta_m$ specifies the magnetic losses in the small-signal range. Based on the impedance $\underline{Z}$ (see (6.5)), the loss factor of the core in conjunction with the complex permeability $\underline{\mu}$ is defined as:

$$\tan\delta_m \overset{\text{def}}{=} \frac{\left|\Im\{\underline{\mu}\}\right|}{\Re\{\underline{\mu}\}} = \frac{\mu''}{\mu'} = \frac{R_S}{\omega L} \tag{6.11}$$

where $R_S$ denotes the series loss resistance.

The magnetic losses, which are covered by (6.11) and cause the phase shift, can be split up into three components:

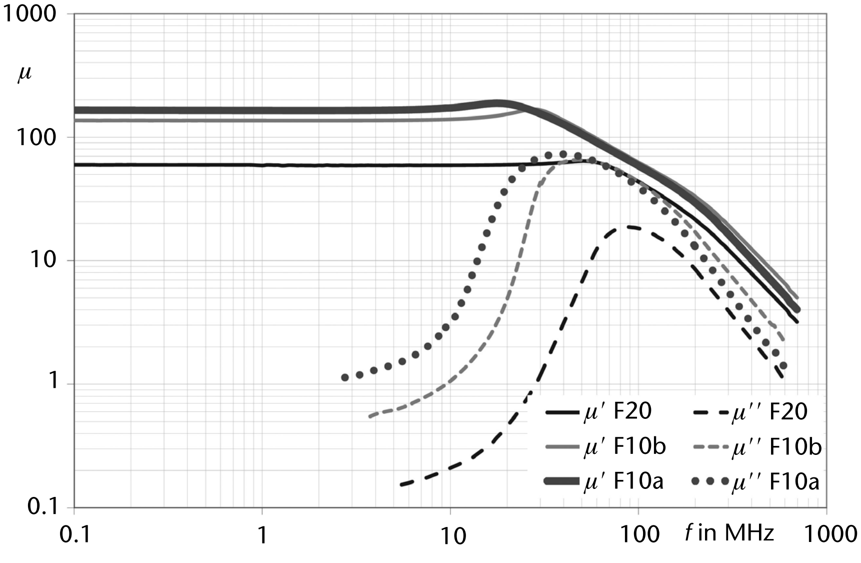


**Figure 6.2** Components of the complex permeability over frequency for a few soft ferrites. (Source: Neosid Catalogue [5].)

- Hysteresis losses ($\delta_h$);
- Eddy (or Foucault) current losses ($\delta_f$);
- Residual losses ($\delta_r$).

This gives the following formula (if tan $\delta_m \leq 0.1$):

$$\tan\delta_m = \tan\delta_h + \tan\delta_f + \tan\delta_r \tag{6.12}$$

Hysteresis losses vanish at very low field strength. Eddy current losses increase with frequency and are negligible at very low frequency. The remaining losses part is called residual loss (see Figure 6.3).

*8. Relative loss factor:* In gapped cores, the material loss factor is reduced by the factor $\mu_e/\mu_i$. This results in the relative loss factor (tan $\delta_m$)$_{\text{gapped}}$:

$$\left(\tan\delta_m\right)_{\text{gapped}} = \frac{\tan\delta_m}{\mu_i} \cdot \mu_e \tag{6.13}$$

where $\mu_e$ is effective permeability.

The table of material properties frequently lists the relative loss factor (tan $\delta/\mu_i$). It is determined at $B = 0.25$ mT and $T = 25°$C and offers the possibility to compare ferrite materials.

*9. Hysteresis material constant:* In some applications, the user cannot be content with small-signal range features. If the field is increased, the hysteresis loop begins to open and the flux density in the core is increased. The hysteresis losses are more noticeable, as they can rise sharply for different $B$-ranges at different frequencies (see Figure 6.3). Their contribution to the total losses can be obtained by means of two measurements, usually at the flux density levels of 1.5 mT and 3 mT and at the working frequency. The hysteresis constant is found from:

$$\eta_B = \frac{\Delta\tan\delta_h}{\mu_e \cdot \Delta B} \tag{6.14}$$

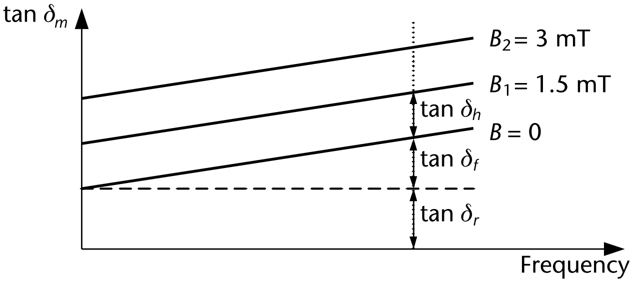


**Figure 6.3** Magnetic losses (tan$\delta_m$) as function of frequency.

where $\Delta\tan\delta_h$ is the difference of the measured loss values ($\tan\delta_m$) at 25°C, corresponding to the above-mentioned B-levels, $\Delta B = 1.5$ mT, and $\mu_e$ is the effective permeability.

The hysteresis loss constant characterizes the material-specific hysteresis losses and is quantity independent of the air gap in a magnetic circuit. If the hysteresis material constant is given in the ferrite specification, the hysteresis loss factor for a certain flux density $B_y$ can be calculated based on the IEC specification:

$$\frac{\tan\delta_h}{\mu_e} = \eta_B \cdot B_y \tag{6.15}$$

*10. Inductance factor ($A_L$):* To make the calculation of a coil inductance easier, the classical inductance factor, known as the $A_L$ value, is given in almost each data sheet (unit: henry · turns$^{-2}$). This value is calculated under standard conditions (10 kHz, 0.25 mT, 25°C) using the core factor and the effective permeability:

$$A_L = \frac{\mu_0 \cdot \mu_e}{\sum\left(\frac{l}{A}\right)} \tag{6.16}$$

For the already specified $A_L$ value, the inductance of a coil using this core and having a number $n$ of turns results:

$$L = n^2 \cdot A_L \tag{6.17}$$

*11. Resistivity ($\rho$):* Ferrite is basically a semiconductor with a DC resistivity in the crystallites varying in a large range: from $10^{-3}$ Ωm for the MnZn type ferrites up to 30 Ωm for the NiZn types. Since the organic adhesives provide an isolating layer between the crystals, the resistivity at the grain boundaries (bulk resistivity) is much higher: 0.1 to 10 Ωm for MnZn ferrites and $10^4$ to $10^6$ Ωm for NiZn ferrites than in the grain interior.

The resistivity is usually specified for DC conditions. It corresponds to the measured resistance $R_{CC}$ of a ferrite segment with defined geometry (length $l$ and cross section $A$):

$$\rho = R_{CC}\frac{A}{l} \tag{6.18}$$

The ferrite as semiconductor has a resistivity, which strongly depends on the temperature $T$. The temperature coefficient:

$$TC\_\rho = \frac{\Delta\rho}{\Delta T} \tag{6.19}$$

basically has a negative value and is material dependent, with a typical value of −0.05 Ωm/°C for MnZn and $-10^5$ Ωm/°C for NiZn.

The resistivity as a function of frequency also has great importance for the behavior of the ISE. At higher frequencies, the crystal boundaries are more or less short-circuited by their capacitance and the resistivity decreases (above 1 MHz, a typical decreasing ratio is one decade of value per one decade of frequency).

*12. Permittivity:* The basic permittivity of MnZn and NiZn is of the order of 10 ($\varepsilon_r = 12$ in the ferrite default models of Maxwell, see Figure 4.15). However, if the bulk permittivity of a ferrite is measured, very different values of apparent permittivity result. This is caused by the conductivity inside the crystallites. The resulting leaky capacitors imply a strong frequency dependence of the permittivity. If the frequency increases, the permittivity goes down (about 0.5 decade of the value over one decade of the frequency).

*13. Temperature coefficient of permeability and temperature factor:* The ferrite permeability is a function of temperature. It generally increases with temperature to a maximum value, the secondary permeability maximum (SPM), and then drops sharply to the value of one. The temperature at which this happens is called Curie temperature ($T_C$). Minimum losses occur at the SPM temperature.

The temperature behavior is described by the temperature coefficient, which has more definitions in ferrite data sheets:

- *Absolute temperature coefficient* that is proportional to permeability rise per °C:

$$TC = \frac{(\mu_i)_{T2} - (\mu_i)_{T1}}{T_2 - T_1} \tag{6.20}$$

where $T_1$ and $T_2$ are two supporting temperatures and $\mu_i$ values in the equation denominator are the corresponding magnitudes of the initial permeability.

This equation defines an average gradient of a straight line between $T_1$ and $T_2$ and gives a good approximation if the curve $\mu_i$ as function of $T$ is linear.

Figure 6.4 shows the initial permeability of a ferrite (type: K2006, manufactured by Kaschke [4]), which has a Curie temperature $T_C > 250°\text{C}$ and is grown for high temperature applications. It is a special ferrite having a greatly extended temperature range. However, its temperature behavior in the most used temperature range (below 100°C) is quasilinear and can be easily modeled.

*Example 6.1* To extend the ferrite model of Maxwell (see Figure 4.15) by implementing the temperature behavior in order to perform simulation at different temperatures, the author simply drew a straight trend line through the reference points with coordinates $\mu_i = 1000$ at $T = -25°\text{C}$ and $\mu_i \approx 3000$ at $T = +100°\text{C}$. The resulting slope of the line is:

$$TC \approx \frac{3{,}000 - 1{,}000}{75 - (-25)} = 20 \left[\frac{1}{°\text{C}}\right] \tag{6.21}$$

Accordingly, the rise edge of the graph in Figure 6.4 can satisfactorily be approximated by:

$$\mu_i \approx 1{,}700 + 20 \cdot (T - 22) \tag{6.22}$$

This procedure significantly enhances the CAE field simulation tools. Usually, the Maxwell material models do not imply temperature behaviors but they can be extended by this term. Returning to Figure 4.15 and looking at the thermal modifier, an extension with (6.22) can be seen in the permeability data line. The author frequently exercised this method and option. The simulation results were close to the reality.

- *Related temperature coefficient* is deduced from (6.20) and relates to the permeability $\mu_{i@RT}$ at the room temperature:

$$TC_{\mu_i} = \frac{1}{\mu_{i@RT}} \cdot \frac{(\mu_i)_{T2} - (\mu_i)_{T1}}{T_2 - T_1} \tag{6.23}$$

For a gapped magnetic circuit, the influence of the permeability temperature dependence is reduced by the factor $\mu_e/\mu_i$:

$$TC_{\text{gap}} = \frac{\mu_e}{(\mu_i)_{T1}} \cdot \frac{1}{(\mu_i)_{T1}} \cdot \frac{(\mu_i)_{T2} - (\mu_i)_{T1}}{T_2 - T_1} \tag{6.24}$$

where $\mu_e$ is the effective permeability.

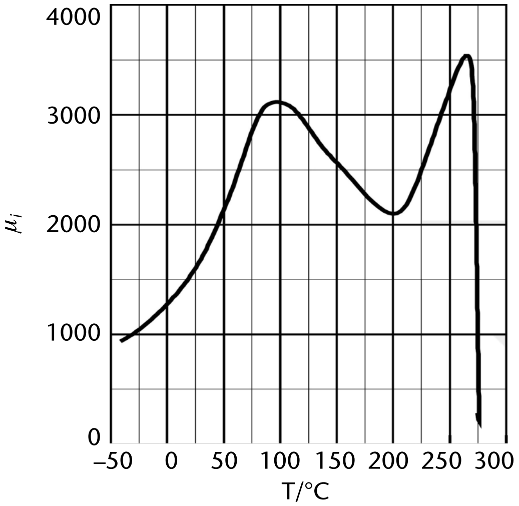


**Figure 6.4** Temperature behavior of the initial permeability of the K2006 ferrite. (Source: Kaschke Components Data Book [4].)

- *Temperature factor* $\alpha_F$ of ferrite is normalized to the unit permeability, expressed in ppm/°C (part per million/°C) and given for a specified temperature range (usually 25°C to 55°C):

$$\alpha_F = \frac{1}{\left[(\mu_i)_{T1}\right]^2} \cdot \frac{(\mu_i)_{T2} - (\mu_i)_{T1}}{T_2 - T_1} \tag{6.25}$$

or, to be more precise, if the change in permeability over the specified temperature range is rather large:

$$\alpha_F = \frac{1}{(\mu_i)_{T1} \cdot (\mu_i)_{T2}} \cdot \frac{(\mu_i)_{T2} - (\mu_i)_{T1}}{T_2 - T_1} \tag{6.26}$$

- *Temperature dependence of the flux density saturation* is important since $B_s$ drops monotonically with temperature and at Curie temperature has fallen to $B_s = 0\text{T}$.

*14. Curie temperature:* The Curie temperature ($T_C$), or Curie point, is the temperature at which certain materials lose their permanent magnetic properties, to be replaced by induced magnetism.

The magnetic moments in a ferromagnetic material are ordered and of the same magnitude in the absence of an applied magnetic field. The magnetic moments in a paramagnetic material are disordered in the absence of an applied magnetic field and ordered in the presence of an applied magnetic field. The transition from ferromagnetic to paramagnetic (or vice versa) occurs at the Curie temperature.

The MnZn soft ferrites usually have Curie temperatures between 100°C and 200°C. Special types with higher values of $T_C$ are manufactured for extended temperature range applications. Increasing the ferrite temperature over the $T_C$, the permeability steeply reaches the value $\mu_i = 1$. The phenomenon is reversible. When the temperature is reduced below $T_C$ again, the magnetic properties are restored.

Some manufacturers are very precise defining $T_C$ as the cross of a straight line passing the levels 80% and 20% of the falling edge of the curve $\mu_i$ versus temperature with the temperature axis (see Figure 6.4).

*15. Time stability, disaccommodation coefficient (d) and factor ($D_F$):* Ferrimagnetic states of equilibrium can be influenced by several troubles. When a soft magnetic is given a magnetic, mechanical or thermal disturbance, the permeability suddenly rises and then decreases slowly with time to the original state. For a defined time interval, this disaccommodation $D$ can be expressed as:

$$D = \frac{\mu_{i1} - \mu_{i2}}{\mu_{i1}} \tag{6.27}$$

where $\mu_{i1}$ and $\mu_{i2}$ are the initial permeability on time $t_1$ and $t_2$ ($t_2 > t_1$).

The permeability appears to have a logarithmic decrease during the time. For this reason, IEC defines a disaccommodation coefficient:

$$d = \frac{\mu_{i1} - \mu_{i2}}{\mu_{i1} \cdot \log\left(\frac{t_2}{t_1}\right)} \tag{6.28}$$

where $t_1$ and $t_2$ are time intervals after the disturbance ($t_2 > t_1$).

As with temperature dependence, the influence of disaccommodation on the inductance change of coil will be reduced by $\mu_e/\mu_i$. Therefore, a disaccommodation factor $D_F$ is defined:

$$D_F = \frac{d}{\mu_{i1}} = \frac{\mu_{i1} - \mu_{i2}}{\mu_{i1}^2 \cdot \log\left(\frac{t_2}{t_1}\right)} \tag{6.29}$$

Usually, ferrite cores are magnetically conditioned by means of an alternating saturation field, which is gradually reduced to zero. Measurements for specified values in data sheets are performed 10 to 100 minutes after this disturbance.

*16. Mass Density (specific weight)* is classically defined as the weight of a ferrite part divided by its volume.

### 6.1.2 Overview of MnZn Ferrite Specifications: Cross-Reference List of Available MnZn Ferrites

To give a general overview of typical values of the soft ferrites provided by different well-known ferrite manufacturers, Table 6.1 lists the main attributes and the extracted values from the corresponding data books (empty records due to missing specification). In addition to the four manufacturers included in the references list, the company Kolektor Magma was considered. Despite a considerable number of empty positions, the table as a functional entity permits to estimate general values of all attributes.

In addition, in Table 6.2, the author provides a helping benchmarking, which compares typical values from Table 6.1 with the default values stated in the material library of Maxwell field simulation software (see Sections 4.3.3 and 4.3.4).

### 6.1.3 Ferrite Core Manufacturing Process: Technical Core Types

A typical flow of the metallurgy process of soft magnetic materials consists of the following production steps [2]:

1. *Raw material availability:* oxides or carbonates of the constituent components with necessary purity are procured;
2. *Incoming inspection:* analysis of chemical compositions;

**Table 6.1** Short Summary of the Main Attributes of the MnZn Soft Ferrites

| *Manufacturer and Ferrite Code* | | *Kaschke* | *Neosid* | *MMG* | *Epcos* | *KM* |
|---|---|---|---|---|---|---|
| *Ferrite Attribute* | *Unit* | *K2006 [4]* | *F08 [5]* | *P11 [6]* | *N22 [7]* | *16G* |
| Saturation flux density ($B_s$) | T | 0.52 | 0.40 | 0.38 | 0.26 | 0.39 |
| at $H$ = | kA/m | 1.2 | 5.0 | 0.8 | 1.2 | 3.0 |
| Remanent flux density ($B_r$) | T | 0.25 | | 0.25 | | |
| Coercive field strength ($H_c$) | A/m | 20 | | | 14 | 20 |
| Initial permeability ($\mu_i$) | | 1700 | 700 | 2250 | 2300 | 2200 |
| Relative loss factor ($\tan\delta/\mu_i$) | $10^{-6}$ | | 20 | 5.0 | <20 | <3.5 |
| at $f$ = | MHz | | 0.8 | | | 0.1 |
| Hysteresis material constant ($\eta_B$) | $10^{-3}$ | | 2.5 | | <1.4 | <0.65 |
| Resistivity ($\rho$) | Ωm | 1 | 1 | 1 | 1 | 3 |
| Curie temperature ($T_c$) | °C | 270 | 260 | 150 | >145 | >150 |
| Temperature factor ($\alpha_F$) | | | | | | |
| for: −30 to +20°C | $10^{-6}$/°C | 6 | 4 | | | |
| for: +20 to +70°C | $10^{-6}$/°C | 6 | 6,6 | | | |
| for: +25 to +55°C | $10^{-6}$/°C | | | 1.5 | 0.9 | 1 |
| for: +5 to +55°C | $10^{-6}$/°C | | | | | 1 |
| for: ·25 to +55°C | $10^{-6}$/°C | | | | | 1.8 |
| Disaccommodation factor ($D_F$) | $10^{-6}$/°C | | 37 | 4 | 4 | <3 |
| Mass Density | Kg/m$^3$ | 4850 | | | 4700 | |
| Frequency range for high Q-values | MHz | | 0.02 to 1.5 | | <0.2 | |

3. *Proportions of the composition:* these are assured by weighting the base materials into the correct proportions.
4. *Mixing of the powders:* this is used to obtain a uniform distribution of the components.
5. *Pre-sintering:* the mixed oxides are calcined at approximately 1,000°C. A solid-state reaction takes place between the constituents and the ferrite is already formed.
6. *Milling and granulation:* these are performed to get specific particle size. A small portion of organic binder is added, and then the slurry is spray-dried to form granules suitable for the forming process.
7. *Forming:* this usually generates the most ferrite part by pressing. The granules are poured into the suitable die and then are compressed. The organic binder acts as an adhesive and a very fragile, so-called green part is formed. Long parts as rods or tubes are extruded through a suitable orifice.
8. *Sintering:* this is required to obtain the final ferrite properties. The green cores are loaded on refractory plates and sintered at temperatures between 1,100 and 1,300°C. That happens in tunnel kilns having controlled temperatures

**Table 6.2** Author's Proposal for a Benchmarking Between Typical Features of MnZn Soft Ferrites and Maxwell Model Default Parameters (Compare with Figure 4.15)

| *Ferrite Default Model of Maxwell Field Simulator* | | | *Typical Parameters Values from Manufacturers Specifications* |
|---|---|---|---|
| *Maxwell Name* | *Units* | *Maxwell Value* | |
| Relative permittivity ($\varepsilon_r$) | | 12 | 10 |
| Relative permeability ($\mu_r$) | | 1000 | 700 to 2,300 |
| Bulk conductivity ($\sigma$) | S/m | 0.01 | 1 to 3 |
| Dielectric loss tangent | | 0 | |
| Magnetic loss tangent | | 0 | |
| Magnetic coercivity ($h_c$) | A/m | 0 | |
| Thermal conductivity | W/m·C | 4 | |
| Magnetic saturation ($b_s$) | Tesla | 0 | 0.26 to 0.52 |
| Core loss model | | None | |
| Mass density | kg/m$^3$ | 4,600 | 4,700 to 4,900 |
| Composition | | Solid | |
| Specific heat | J/kg·C | 750 | |
| Young's modulus | N/m$^2$ | $1.19 \times 10^{11}$ | |
| Poisson's ratio | | 0 | |
| Thermal expansion coefficient | 1/C | $1.0 \times 10^5$ | |

and atmosphere distribution. An undesired linear shrinkage of up to 20% takes place and must be preliminarily considered;

9. *Sinter check:* testing of electrical and mechanical properties;
10. *Finishing* is the last fabrication step, which requires grinding and lapping operations to provide the fulfilment of the part dimensions and tolerances;
11. *Final step:* final inspection, delivery release, Outgoing inspection, and sales.

Apart from structural core shapes that are customer application specific, the manufacturers offer a large spectrum of conventional core types. There exist a great number of technical core types (see Figure 6.5) and an important number of standards stating shapes, dimensions, variants and other important features to provide cores interchangeability between different providers.

The author tried to summarize and to arrange the vast volume of information. The following classification reflects these efforts. Good references for additional knowledge, including material data and selection guides, are [6] and [7]. Moreover, the reference [7] contains a helpful list with the IEC standards regarding ferrites, ferrite cores, testing and measuring procedures and terms (IEC 60424, IEC 62317, IEC 61631, etc.).

To facilitate the interchangeability, the provider himself offers a cross-reference list as the example in Figure 6.6.

The proposed author's classification of the typical core types consists of six main classes with subclasses, as follows:

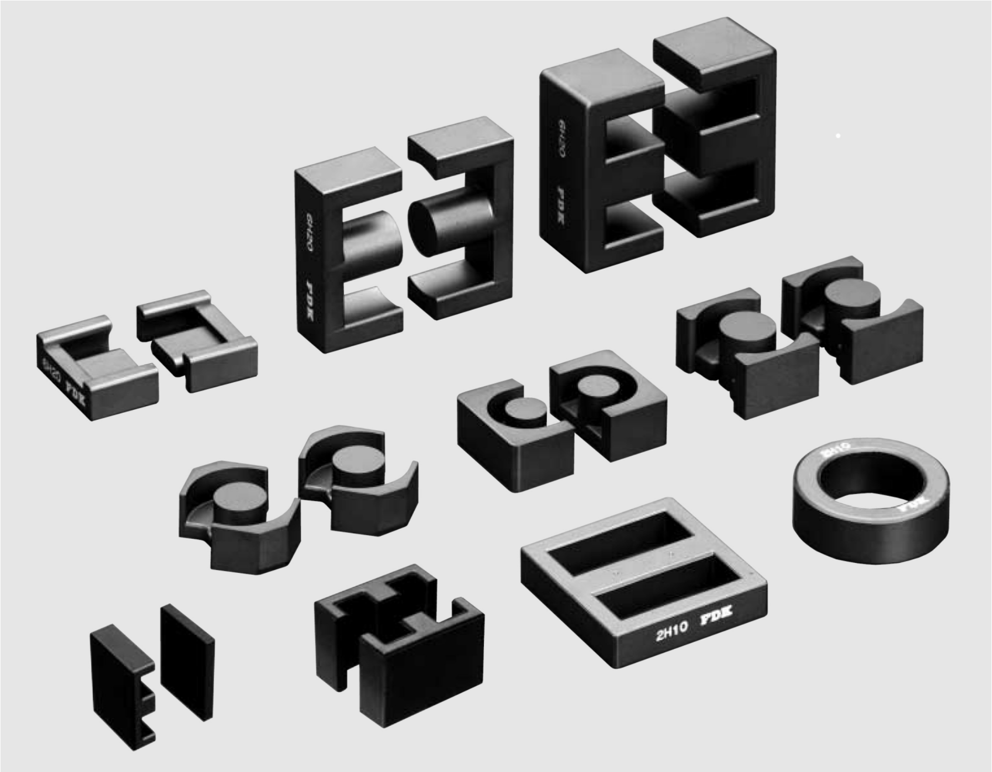


**Figure 6.5** Three-dimensional pictures of soft ferrite cores. (Source: FDK-Japan Corporation Data Book)

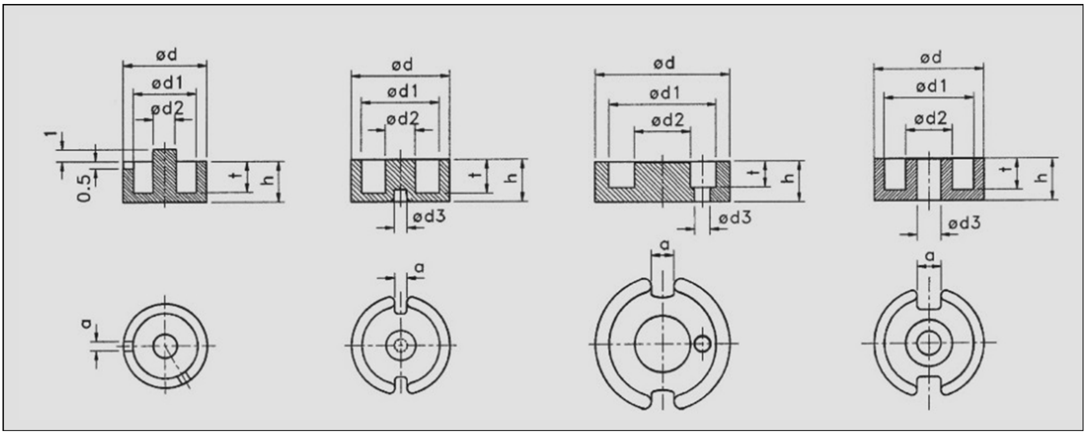
1. *Round pot* cores (*P*-type) complying with IEC standards:
   - *Half-pot* types for single usage and implementation in IPS [8, 9] (see the first line in Table 6.3; the fourth version in this line is implied in Examples 4.1 and 5.2);
   - *Pair of pot* types for face-to-face mounting (transformer) or back-to-back mounting (double coil system for sensing and compensation);
   - *PM*, *PQ*, and *RM* types that are special designs and offering high density of flat mountings on board.
2. *E-shaped* and *E-derived* cores (see the second line in Table 6.3):
   - *E* type is the classical shape with a large diversity of dimensions (single or double *E*);
   - *EER and EF/EFD* type with cylindrical shape of the leg (winding-advantageous);
   - *EI* type, which is a pair consisting of an *E* core and an *I* core;
   - *EP/EPO*, *EPX*, *ER* and *EV* types;
   - *ETD* type known as economic transformer design.
3. Frame cores and planar cores:
   - *FU and FUR* types that have frame shapes and can perform the highest closed magnetic circuits;
   - *PE* type, which is an *E* core with flat, planar geometry;
   - *Low profile* types (*ELP*, *EQ*, and *RM*).

| Application | Permeability | Kaschke | Epcos | Ferroxcube | Sumida | Tridelta | Kolektor | MMG | TDK | ACME | TDG | DMGECC | Samwha | Cosmo | Magnetics |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | < 1000 | **K 300**<br>**K 600** | (M33) | 3D3<br>3B1 | Fi262 | MF143 | 10G | | | | | | SM-8T | | |
| power ferrites for high frequencies | 700 - 1300 | | N 59 | | | | | | | P6 | TP5B | DMR51 | PF-L2 | | L |
| | 1300 - 1500 | **K 2001** | **N 49** | **3F35** | Fi327 | MF114 | **76G**<br>75G | | PC50 | **P51** | **TP5** | **DMR50** | **PF-L1** | | |
| power ferrites with low losses | 2000 - 2500 | **K 2004** | **N 27** | **(3C80)**<br>**(3B8)** | Fi324 | MF196 | | **F5** | | | | | | | |
| | | **K 2006** | **N 67** | | | MF198<br>**MF198A** | **15G** | **F44** | | | | | PL-5 | CF196 | R |
| | | **K 2008** | **N 87** | **3C90**<br>3C94<br>3F3 | Fi325<br>Fi335 | **MF102** | **35G**<br>45G | **F44A**<br>**F48** | PC40 | **P4** | **TP4** | **DMR40** | PL-7<br>**PL-9** | | **P** |
| | | **K 2010** | N 97 | 3C94<br>**3C96** | | MF106 | 65G | | PC44<br>PC47 | P41 | TP4A<br>TP4S | DMR44<br>DMR47 | PL-15 | **CF297**<br>CF139 | |
| power ferrites with high flux density | 1500 - 2000 | **K 2024** | **N92** | **3C30**<br>3C92 | Fi328<br>**Fi329** | MF104 | | | PE22 | P42 | **TP4F**<br>TP4G | **DMR24** | **PL-HB** | | |
| | | **K 2026** | N 93 | **3C92** | | MF108 | 85G | | PE90 | **P49**<br>P48 | TPB22 | | PL-HB | | |
| high perm power ferrites | 2500 - 4000 | **K 2500** | N 72<br>N 51<br>N 95<br>N 96<br>N 45 | **3C81**<br>3C91<br>3C95<br>3C97 | Fi395 | **MF95** | | **F5A**<br>F5C<br>F63 | PC95 | P45<br>P46<br>P47 | TP4B<br>TP4D<br>TP4W | | ST-30B | **CF124**<br>CF130 | F |
| Filter ferrites | 2000 - 3000 | **K 2005** | N 22<br>N 48 | **3H3**<br>3B7 | | MF183 | 26G | **P11**<br>P12 | | N4 | TH2 | DMR70 | SM-23T | **CF140** | |
| Broadband and EMC ferrites | 3000 - 4500 | **K 4000** | T 57<br>**N 30** | **3S1**<br>**3C11**<br>3E28<br>3E4 | **Fi340** | **MF103**<br>**MF193** | **19G** | F9 | DNW45 | A043 | **TH4** | **R4K** | SM-43T | | |
| | 4500 - 5500 | **K 5500** | **T 65** | | | | | F9C | HP5 | **A05** | TS5 | R5K | SM-50 | CF255<br>CF195A | J |
| | 5500 - 8000 | **K 6000**<br>**K 7000**<br>**K 8000** | **T 35**<br>T 36<br>T 37 | **3E25**<br>3E26 | | **MF197** | **22GN**<br>22G<br>42G | **F10** | HS72 | A07<br>N07 | TS7 | R7K | **SM-60**<br>SM-70S | **CF190**<br>CF275<br>CF197 | |
| medium perm ferrites with high $T_C$ | 5500 - 8000 | **K 6001** | | **3E27** | Fi360 | | 23G | | | **A061**<br>**A062** | | | | | |
| high perm ferrites | 8000 - 12000 | **K 10000**<br>**K 12000** | **T 38** | **3E5**<br>3E55<br>3E6 | **Fi410**<br>**Fi412** | MF199 | **12G**<br>32G | **F39**<br>FTA | **HS10**<br>H5C2<br>HS12 | **A10**<br>A102<br>A121 | TH10<br>TS10 | **R10K**<br>R12K | **SM-100**<br>SM-120 | **CF199** | W |
| | > 12000 | **K 15000** | T 66<br>**T 46** | **3E7**<br>3E8<br>3E9 | **Fi415** | | **52G** | | H5C3 | **A151** | **TL15**<br>TL13 | **R15K** | SM-150 | | H |

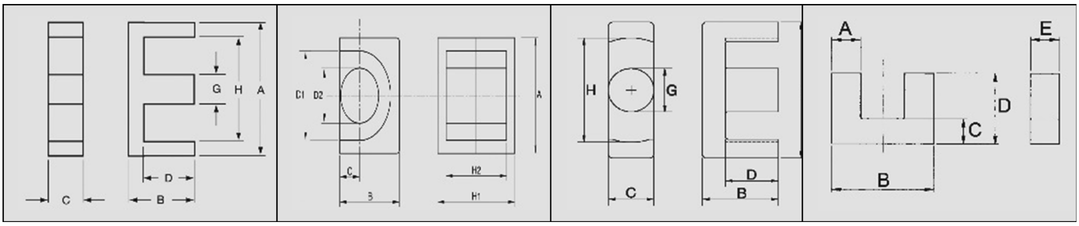
**Figure 6.6** Cross-reference list for MnZn ferrites offered by a number of ferrite manufacturers. (Source: Kaschke Components Data Book [4].)

**Table 6.3** Typical Structural Forms of Soft Territe Cores for Inductors (Source: Neosid Catalogue [5])

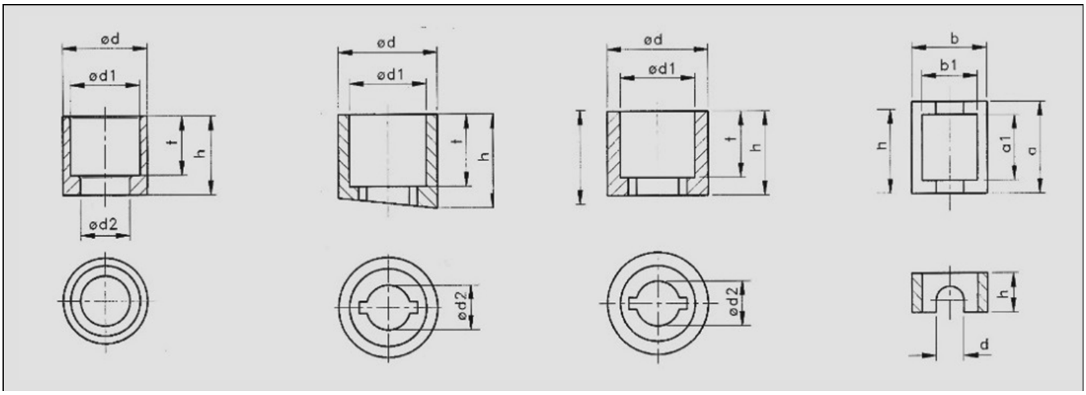
*Pot Cores*


ød
ød1
ød2
ød3
1
0.5
t
h
a

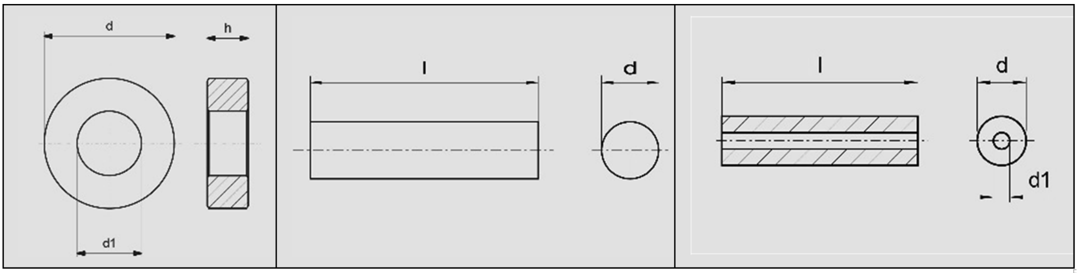
*E- / EP- / ETD- / U-Cores*


G
H
A
C
D
B
C1
D2
H2
H1
E

*Cup Cores*


ød
ød1
ød2
t
h
b
b1
a1
a
d

*Ring- / Rod- / Tube-Cores*



4. *Cup-derived* cores (see the third line in Table 6.3);
5. *Geometrical body* cores (see the fourth line in Table 6.3):
   - *Role* core;
   - *Rod* core with versions: solid rod, solid flat-sided, solid fluted rod;
   - *Tube* core with versions: hollow rod, hollow fluted rod;
   - *Ring* core (toroid);
   - *U* and *UI* cores.
6. Smart core designs for SMT.

In addition to magnetic and electrical properties (Section 6.1.1), the ferrite technical cores are characterized by mechanical properties:

- *Mechanical requirements during assembling* have to be fulfilled. Basically, ferrites are ceramic materials and thus are brittle or sensitive to any shock, fast changing, or tensile load. Especially high cooling rates under ultrasonic cleaning and high static or cyclic loads can cause cracks or failures of the core, that dramatically change the magnetic core coil features. There are two modes of crack growth: fast (critical) or slow (subcritical) crack propagation. In the first case, spontaneous breakdown occurs. In the second case, the crack propagates slowly and the sample can only fail if a critical length is achieved.
- *Mechanical stress limit* defines a threshold pressure value, above which the magnetic properties (first the permeability) dramatically alter. Embedding the ferrite core (e.g., in plastic) can induce these stresses.
- *Magnetostriction* is defined as the relative change in length of the ferrite core under the influence of a magnetic field. The greatest relative change occurs at saturation magnetization. The effect is significant when the core is operated at frequencies below 20 kHz and then undesired audible frequency phenomena (distortion, etc.) occurs.
- *Resistance to radiation* specifies a radiation limit without relevant variation of the core coil inductance (<1% is usually accepted).

In parallel to ferrite cores, a couple of manufacturers use another material to produce technical cores (see Table 6.3), namely the carbonyl iron. It is basically a mixture of iron powder with an adhesive.

It concerns a highly pure iron, prepared by chemical decomposition of purified iron pentacarbonyl $Fe(CO)_5$. The component iron and white damp of the raw material are separated by heat-treatment. Particulate hard material results having a Fe purity of 98%, a carbon-content below 1% and some impurities like oxygen and nitrogen. This grey powder, composed of spherical microparticles, is used to generate cores with small measuring tolerances (compression-molding procedure). A heat post-treatment consisting of white heart tempering of the hardness cores can give them a soft character, increasing the Fe purity up to 99.6% and reducing the carbon-content to 0.1% [4].

The carbonyl iron is a BASF invention (1925) and has a vast field of applications in audio technique, electronics, radar, and others.

The magnetic field distribution in carbonyl iron cores is uniform and the hysteresis losses are very low. These positive features are fulfilled over a large frequency range. The Q-factor reaches excellent values in a frequency range between 50 kHz and 200 MHz. The material has a long-term stability and gut performance across a wide range of temperatures.

But for all that, the material has a high ruggedness against disturbances provoked by environmental magnetic fields. This property makes the carbonyl iron cores suitable for inductive sensing elements able to work in electrical welding plants. The saturation flux density achieves high levels. The clearly negative parameter is the relative low values of the initial permeability ($\mu_i < 70$). The Curie temperature lies in the range 120°C to 150°C [4].

## 6.2 Permaloy and Mu-Metals

Ferrites can theoretically be formed as thin plates. Nevertheless, there is practical limitation of the thickness. Making very thin foils is process-related and therefore is not possible.

Foils with magnetically soft properties are required for some ISEs. Besides the classical application of these foils as electromagnetic shields, they perform two important tasks:

1. Play the three specific roles of a conventional coil core (see Section 5.3). The coils, which are now considered, are not wired inductors but rather printed coils on an insulator carrier substrate (see Section 5.4).
2. The foil attached on the substrate reverse can be punctually driven into magnetic saturation by an external permanent magnet. This local change has an influence on the inductance of the planar coil.

The spectrum of suitable materials naturally debuts with the permalloy (permeable alloy). It is a nickel-iron soft magnetic alloy, with about 80% nickel and 20% iron content. Permalloy, invented in 1914 by physicist Gustav Elmen at Bell Telephone Laboratories (78.5% nickel and 21.5% iron), has exciting history and applications. It was initially developed and produced for the inductive compensation of telegraph submarine cables. The method declined in the 1930s, but Permalloy found many other applications in the electronic industry.

The alloy is notable for its properties:

- Very high magnetic permeability ($\mu_r = 100{,}000$ typically), compared to several thousand for steel;
- Reasonable saturation flux density ($B_s < 1\text{T}$);
- Low coercive field strength ($H_c$);
- Low electrical resistivity ($\rho = 10\text{E-}6\ \Omega\text{m}$);
- Near zero magnetostriction;
- Suitable for magnetic shielding to block low frequency magnetic fields.

The disadvantage of permalloy is that it is not very ductile or workable. In addition, this alloy was protected by a Western Electric patent. To avoid the monopoly granted by this patent and to improve the material properties (better plastic deformation), Telcon Ltd./UK invented a new material called Mu-metal, which additionally comprises 5% copper and 3% chrome.

After cooling, the melted Mu-metal alloy can be processed as sheet and foil with a thickness between 0.1 and 5.0 mm or wire. The green material can furthermore be stamped out, etched, bent, soldered, or galvanic coated. It has a much higher permeability (up to 300,000). After a mechanical processing, the permeability drops dramatically off and the material must be unconditionally annealed (at 1,000 to 1,200°C, in vacuum, hydrogen, or inert gas).

Metal Magnetic Shielding manufactures a magnetically annealed Mu-metal named ASTM A753 Alloy 4 with very high features [10]. It is a soft ferromagnetic alloy with 80% nickel, 4.5% molybdenum, and balanced iron.

It offers very high permittivity (up to 500,000) with minimum hysteresis loss, high saturation flux density (<0.76T) and very low coercive field strength (<0.6 A/m). The high permeability provides a low reluctance path for magnetic flux making the material usable in magnetic shields against DC and AC fields. To reach these magnetic performances, the alloy needs a full final treatment in a dry hydrogen atmosphere furnace at 1100°C to 1180°C for 2 to 4 hours. The annealing process increases the grain size and hence the magnetic domains. However, these enlarged grains are susceptible to severe shocks or mechanical working and the permeability will decrease in these areas. This can be rectified by a re-anneal.

Later on in 1940, Bell Labs invented the molybdenum permalloy (81% nickel, 17% iron, and 2% molybdenum) and finally the supermalloy (79% nickel, 16% iron, and 5% molybdenum) that have much better performances: high permeability and low coercivity.

At the end of this technological improving competition, the permalloy mixture was improved, so that today both alloys, permalloy and Mu-metal, possess similar features.

A leading global manufacturer of advanced magnetic materials and related products, namely Vacuumschmelze (VAC)/Hanau-Germany offers a large spectrum of such materials. To simplify the selection process, VAC provides in the product brochure [11] a classification of the soft magnetic materials into nine groups. Each group is briefly described from features up to applications. The author tried to continue this effort and compacted them in tabulation (Table 6.4).

Besides various possibilities, VAC offers the tuning option of the BH curves. The three shapes that can be fit in the alloy are shown in Figure 6.7.

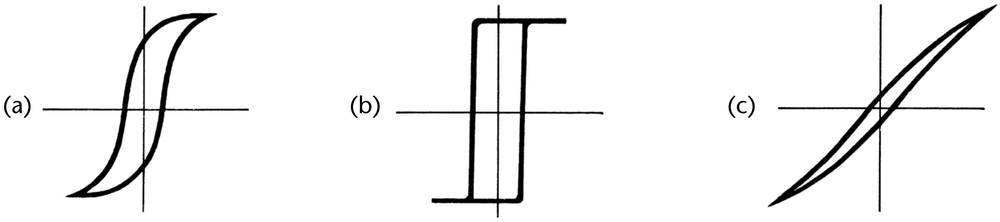


**Figure 6.7** Example of possible hysteresis loop shapes of VAC alloys: (a) round (R), (b) rectangular (Z), and (c) flat (F). (Source: Vacuumschmelze (VAC) [11].)

**Table 6.4** Summarizing Overview of the Various Magnetic Soft Alloys Provided by Vacuumschmelze

| *VAC Magnetic Material-Group* | *Alloy Name* | *Features, Remarks* | *Hysteresis Loop* | | |
|---|---|---|---|---|---|
| | | | *R* | *Z* | *F* |
| NiFe alloys (72–83% Ni) | MUMETAL<br>VACOPERM<br>ULTRAPERM | High initial permeability,<br>Low coercivity,<br>Relatively high saturation (0.9T) | √ | √ | √ |
| | CRYOPERM | Suitable for low temperatures. | | | |
| | RECOVAC | Alloy with higher hardness and wear resistance,<br>Relatively high saturation (0.8T) | | | |
| NiFe alloys (54–68% Ni) | PERMAX M | High initial permeability,<br>High saturation (1.5T). | √ | | |
| | PERMAX F | Adjustable flat hysteresis loop<br>High pulse permeability | | | √ |
| NiFe alloys (54–68% Ni) | PERMENORM 5000 | High saturation (1.5T–1.6T),<br>Medium initial permeability,<br>Trim-able microstructure and Hysteresis loop. | √ | | |
| | RECOVAC | High wear resistance reached by heat-treatment. | √ | | |
| NiFe alloys (35–40% Ni) | PERMENORM 3601 | Fine grained isotropic structure,<br>High electrical resistivity. | √ | | |
| | MEGAPERM | Fine grained isotropic structure,<br>High electrical resistivity,<br>High saturation. | √ | | |
| | CHRONOPERM | NiFeCr-Alloy. | √ | | |
| NiFe alloys (≈ 40% Ni) | THERMOFLUX | Adjustable Curie temperature (typ. 55°C) ⇒ Saturation polarization heavily dependent on the ambient temperature. | √ | | |
| SiFe alloys with 3% Si | TRAFOPERM | Nearly isotropic 3% SiFe,<br>Saturation slightly below that of iron (RFe),<br>Electrical resistivity four times higher ⇒ lower remagnetization losses. | √ | | |
| Magnetic ultrapure iron alloy (99.98%) | VACOFER | Highest purity solely by sintering and absence of pores,<br>Very low coercivity, suitable for yokes with CML (Section 3.3.3),<br>Poorer corrosion resistance,<br>Low electrical resistivity. | √ | | |
| CoFe alloys with 17–50% Co | VACOFLUX | Highest saturation polarisation (up to 2.35T) of known alloys,<br>Low coercivity, very low losses. | √ | | |
| | VACODUR | Further development of VACOFLUX with respect to improved mechanical properties. | √ | | |

*(continues)*

**Table 6.4** *(Continued)*

| *VAC Magnetic Material-Group* | *Alloy Name* | *Features, Remarks* | *Hysteresis Loop* | | |
|---|---|---|---|---|---|
| | | | *R* | *Z* | *F* |
| High-performance amorphous and nano-crystalline alloys | VITROVAC 6025<br>VITROVAC 6030<br>VITROVAC 6150 | Amorphous alloys: 70 to 80% transition metals (Fe, Co, Ni) and 30 to 20% metalloids (B, C, Si),<br>Vitrovacs are Co-base alloys,<br>Despite amorphous state, very high hardness of ribbons,<br>Optionally with heat treatment and glass-like solidification,<br>High permeability (2,000 to 5,000) and very high for F hysteresis loops (up to 60,000),<br>High saturation up to 1T (see Figure 6.9) | | √ | √ |
| | VITROPERM 800 | Fe-base alloys,<br>Crystal grains with diameter in nm-range due to the crystallization heat treatment,<br>Potential matching of peak values: permeability, coercivity and losses<br>Superior in saturation flux density and temperature stability. | | √ | √ |

After: [11].

In terms of the initial magnetization curve (compare with Figure 6.1), Figure 6.8 refers to a couple of established VAC materials summarized in Table 6.4. VAC is also known through the VITROVAC line, whose components are benchmarked in Figure 6.9.

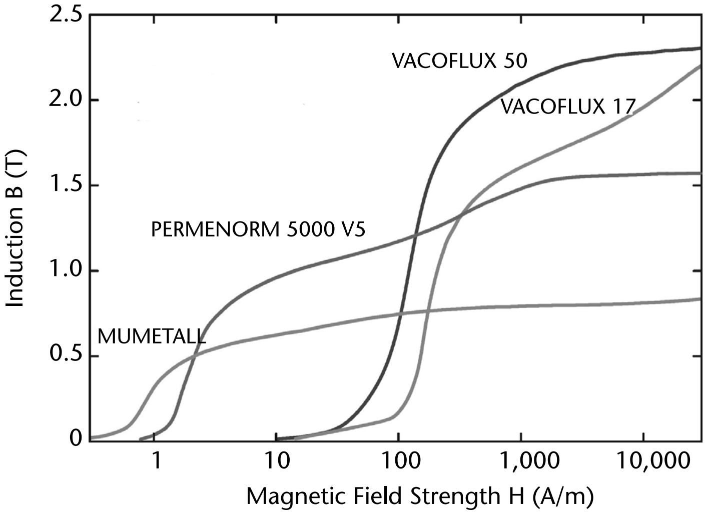


**Figure 6.8** Graphical benchmarking of initial magnetization curves of several VAC alloys. (Source: Vacuumschmelze (VAC) [11]).

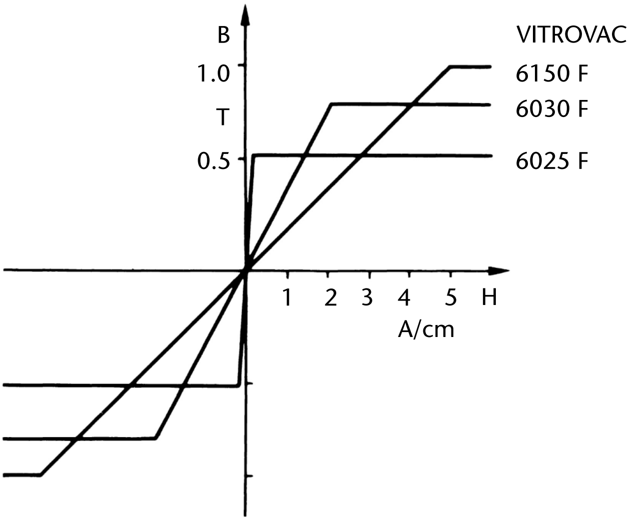


**Figure 6.9** Typical hysteresis loops (semi-schematic representation) of VITROVAC 6025, 6030, and 6150. (Source: Vacuumschmelze (VAC) [11]).

## 6.3 Soft Iron Alloys

The previous magnetic materials (ferrites and permalloy) are essentially adequate for unmovable cores wherein the sensing operation occurs by means of contactless actuation (see Section 1.2.1).

In contrast, inductive sensors with a movable cores that has a mechanical linkage to the target (e.g., the LVDT representation in Figure 3.17) use core materials selected to provide the needed amount of inductance change as it moves into the coils. In addition, this core has to have favorable properties such as hardness, long-term stability, temperature stability, low power loss at the operating frequency, and corrosion resistance.

Traditional materials for movable cores are iron or nickel-iron alloys, optionally with some other materials added to tune the magnetic and mechanical properties.

Several alloys summarized in Table 6.4 are recommended by the manufacturer Vacuumschmelze also for movable cores [11]. These are:

- *VACOFER*, which is a soft magnetic ultrapure iron alloy (99.98%) and is suitable for mechanical processing (sawing, drilling, etc.);
- *ULTRAPERM*, *PERMAX*, *PERMENORM*, *MEGAPERM*, which are nickel-iron alloys with different Ni-contents and corresponding with varied properties (see Table 6.4).

In addition, the reference [12] states two suitable materials, namely:

- *Ni-Span-C* is a trademark of the Special Metals Corporation (www.specialmetals.com). It is an iron-nickel-chromium-titanium composition with controllable thermoelastic coefficient. The temperature dependence of its permeability is also adjustable and can be adjusted to be very low in the normal temperature range for industrial use (–40°C to 85°C).

- *CarTech Hiperco 50B* is a cobalt-iron alloy (nominal composition: 48.50% Co, 0.01% C, 0.20% Mn, 0.05% Si) offered by Carpenter Company (www.cartech.com). It is a high-magnetic-saturation alloy with higher permeability than Ni-Span-C and can be supplied in bar or wire form.

The final annealing process, which relieves the mechanical stress from machining and makes the cores magnetically uniform, is succinctly described in [12]. It should be performed by the client according to the material manufacturer specification. The core is maintained at an elevated temperature for the specified period of time. In the oven, oxygen is excluded to prevent oxidation, and a reducing atmosphere is generally used.

## References

[1] Goldman, A., *Modern Ferrite Technology*, 2nd edition, Springer-Verlag US, 2006.

[2] Chen, C.-W., *Magnetism and Metallurgy of Soft Magnetic materials*, New York: Dover Publications, Inc., 1986

[3] Dorf, R. C. (ed.), *The Electrical Engineering Handbook*, Boca Raton, FL: CRC Press/IEEE Press, 1997.

[4] Kaschke Components GmbH, *Kaschke_Katalog_2016_web_k.pdf*: http://www.kaschke.de, 2016.

[5] Neosid Pemetzrieder GmbH & Co. KG, *Neoside_catalogue_2015_complete*.pdf: http://neosid.de, 2015.

[6] MMG Canada Limited, *Product Line catalogue—Soft Ferrite materials*: http://www.mmgca.com, 2017

[7] EPCOS TDK Group, *Data Book_2013*: http://www.epcos.de, 2013.

[8] IEC-International Electrotechnical Commission, International Standard IEC 62323, *Dimensions of half pot-cores of ferrite for inductive proximity switches*, www.vde-verlag.de/iec-normen, Germany: 2006.

[9] Esguerra, M., *Ferritkerne in induktiven Naeherungsschaltern*, Siemens Matsushita Components GmbH, Components 30, Band 4, 1992.

[10] http://www.mumetal.co.uk/?p=100.

[11] Vacuumschmelze GmbH & Co. KG, *Soft Magnetic Materials and Semi-finished Products*, http://www.vacuumschmelze.com, 2002.

[12] Nyce, D. S., *Linear Position Sensors*, Hoboken, NJ: John Wiley & Sons, 2016.

CHAPTER 7

# Evaluation Electronics of the Inductive Sensors

This chapter begins the third part of the book, which deals with the electronic functional units and stages of the inductive sensors ISs.

The first part (Chapters 1 to 3) dealt with the fundamentals of the inductive sensors and was followed by the second part (Chapters 4 to 6), which thoroughly presented the first element of these sensors, the inductive sensing element (ISE).

This chapter and Chapters 8 to 13 highlight the second, central subassembly of the ISs, which is evaluation electronics (EE) (see Figure 1.2).

The author proposes at first a generic block diagram for this component, which should universally be valid for the very large diversity of ISs. The functional blocks and their composing electronic stages are solely defined showing their functionality. Detailed information about them as variants, operating principles, implementations, and use in IS, and so forth, are available in Chapters 8 to 13.

After a description of this generic block diagram, this chapter continues with a unique presentation of the implementation and integration levels of this functional diagram starting with historical discrete circuitries on transistor level and ending with future-oriented software-defined sensors (SDSs).

## 7.1 Generic Functional Diagram of the Inductive Sensor's Evaluation Electronics

The above-mentioned functional schematic of the technical inner life of an IS is shown in Figure 7.1. It represents a further development and itemization of the top level representation in Figure 1.2.

The inductive sensing element (SE in Figure 1.2) was detailed in Chapters 4 to 6. Now this component is depicted in Figure 7.1 by the entity of three inductors.

The evaluation electronics (EE in Figure 1.2) is presently divided into four basic functional units:

1. Front-end electronics (FEE) marks the low-signal and high-sensitivity part of every sensor. In this part, the FEE usually consists of the following electronic stages:
   a. Oscillator (OSC), which is core unit of each IS and has the task of supplying the ISE. Moreover, for some systems, it is also involved in the evaluation of the physical measurand (see systems in Section 3.3.6). Due to their importance, Chapter 8 is fully dedicated to oscillators.

   b. Local trimming (LTM) provides an adjustment in FEE—basically of the OSC—to set the main IS features related to the sensing range. The popular *teach-in sensor calibration* can manually be made on the sensor assembling line or automatically at the end of the production line (testing and final check) with support of the digital control unit (DCU). Chapter 12 deals with trimming procedures.
   c. Amplifier (AMP) picks up the weak signals coming from ISE (range between a few tens and a few hundreds of millivolts) that are the result of the primary evaluation of the measurand and makes them suitable for the next stage.
   d. Detector (DET) converts the AC signals that originated from the ISE into DC signals, which are easily processed. The large variety of detection solutions—from analog to digital versions—available in ISs are analyzed in Chapter 9.
   e. Temperature compensation (TCP) is a fundamental item of the ISs. Chapters 4 to 6 mention the negative influence of the temperature on the operating mode, the component features, and the materials belonging to the ISE. On the other hand, the specific demands of the industrial applications sensors for stability over temperature in a large temperature range (see Section 1.3.5) are very tough. Possibilities to solve this conflict are presented in Chapter 13.
2. Back-end electronics (BEE) follows the FEE and continues the signal processing from a primary status offered by FEE to the status of sensor final output signals. To maintain the generic character, the structure contains two parallel signal paths, analog and digital, which can coexist. However, typical sensors can contain either.
   a. Signal conditioning (SCD) opens the analog signal path and keeps on doing a signal preconditioning (see Chapter 9).
   b. Linearization (LIN) carries on with the preparation of the required or of the standard compliant sensor characteristic curve (see Sections 1.3.2, 1.3.3, and 2.1.2). The linearity of the IS represents a second main demand besides temperature stability. The number of IS users, which are depicted with a monotonic sensor characteristic curve (Figure 1.3) without high linearity requirements appears rather manageable. The linearization methods are a recurrent theme in the next chapters.
   c. Analog output stage (AOS). The analog signal path ends with this stage, where a final amplification occurs, so that the output signals *V/I-OUT* (voltage or current) are standard-compliant (see Sections 1.3.7.2 and 2.1.2). Chapter 10 deals with specific output stages.
   d. Threshold comparator (CMP) opens the digital signal path and is often a Schmitt trigger [1]. From the functional point of view, this is a regenerative comparator with hysteresis implemented by applying positive feedback to the noninverting input of a comparator or a differential amplifier. The circuit is called a trigger because the output retains its value until the input changes sufficiently to trigger a change at the output. CMP converts the analog signal coming from DET to a digital output signal. The trigger operation is required by the standard demand for an implemented

   differential travel *H* of sensor binary outputs (Table 1.2, Figure 2.3, and Section 2.1.1.3). The trigger threshold and the hysteresis can be fixed or programmed via DCU. The signal threshold discrimination is a theme in Chapter 9.
   e. Combinational logic unit (CLU) symbolizes the logical links through different stages and resulting signals, which provide the sensor-specific logical functions (e.g., the switching functions and switching output types (see Table 1.2 and Figure 1.3.7.1, 2.1.1.4). These various logical operations can be ex-factory hard-coded or are programmable again via DCU.
   f. Switching output driver(s) (SOD) provides the switching output signal(s), which are symbolized by the binary signal line *OUT*. It is a very important stage, which complies with a wide number of standard demands regarding the sensor switching capability, output type: HSD, LSD, push-pull (see Table 1.2), output current efficiency, switching dynamic performances (Section 2.1.1.5), and so forth. Such aspects are a topic of Chapter 10.
   g. Short-circuit protection (SCP) shares a circuitry for both outputs (analog and binary), which continuously monitors the output currents and immediately acts to protect the outputs if the current exceeds a maximal value (over-current) or if a short-circuit occurs at the output (see Section 2.1.1.5, part e).
   h. Display/optical indicator (D/OI) is realized with a lot of possible implementations: binary indicators with light-emitting diodes (LEDs) or liquid crystal display (LCD), analog bargraph indicators, numerical (7-segment) or alpha-numerical (16-segment or 35-dot matrix), as well as display screens. The minimum configuration is given by LED-indicator(s).
3. Power supply and protections (PSPs) are essential functional units of any sensor installed in industrial plants. In Figure 7.1 this block is a placeholder not only for sensor internal supplying but also for sensor internal protection functions, which are mandatory for industrial sensor implementations. The inventory list of these tasks contains:
   a. Power management;
   b. Generation of the sensor internal supply voltage(s): $+V_{CC}$, $+V_{DD}$, $+V_{SS}$ based on the usually singular unipolar supply voltage $+V_B$;
   c. High-voltage protection;
   d. Surge protection against high-energetic impulse voltages (see Section 2.2.4);
   e. ESD protection against electrostatic discharges (see Section 2.2.1);
   f. EFT protection against fast transients (see Section 2.2.3).
4. Digital control unit (DCU) is an advanced functional stage, which belongs to modern, state-of-the-art inductive sensors (see Section 7.4). As a result of sensor specific constraints (low costs, limited free space inside the sensor housing, low current consumption, etc.), the DCU has a moderate extent commonly consisting of
   a. Finite-state machine (FSM), which is defined in the automata theory as a mathematical model of computation and represents the second class of automata after the combinational logic. It is generally an abstract machine that can be in exactly one of a finite number of states (equivalent to the

sensor states) at any given time. The FSM can change from one state to another in response to a change of some inputs or conditions. Such changes are called transitions. An FSM is formally defined by the list of its states, its initial state, and the conditions for each transition.

b. Memory (MEM), which normally consists of a volatile memory to store temporary data and a nonvolatile part to store sensor parameter and/or programs;
c. Clock generator (CLG) principally acts as a clock signal maker for the FSM but also as a time base for generation of specific sensor times (see Section 1.3.6).
d. Programming and control (PGC) is responsible for the execution of several programming subroutines (sensor calibration, teaching of different parameters, etc.).
e. Fieldbus interface (FIF), when up-to-date inductive sensors unconditionally have at least a serial bidirectional communication channel I/O-BUS, which is either a serial interface (see Section 1.3.7.4) or a fieldbus (see Section 1.3.7.5). This theme is the subject of Chapter 14.

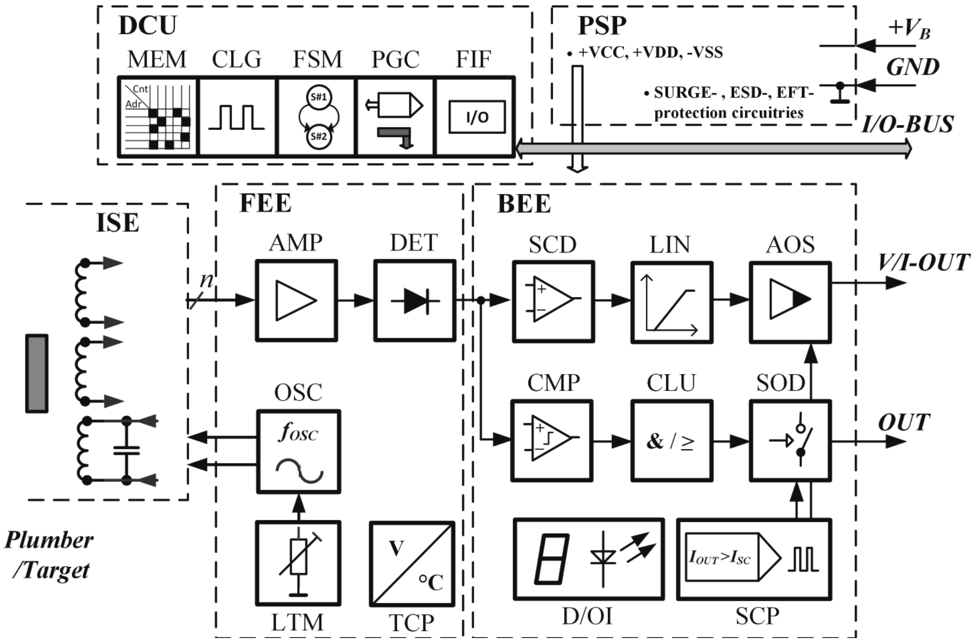
Due to the widespread use of inductive proximity sensors (IPSs) for industrial automation applications in the last 50 years, their evaluation electronics was continuously developed and integrated on larger and larger scales. In addition, the involved silicon foundries, as well as their applicants—namely, sensor manufacturers—openly published these technological steps. Correspondingly, the evolution was open. With this rich volume of information, we can elaborate a morphological map, which shows the evolution steps of the EE of IPS (see Figure 7.2). The representation is but an example and can also be made for other inductive sensor types.

The morphological map in Figure 7.2 starts with the lowest technological level of discrete designs and passes more phases of integration, starting with the IC levels, going through application-specific integrated circuit (ASIC) levels, and ending with the top level of systems-on-chip (SOC). This evolution steps are subject of the Sections 7.2 to 7.5, respectively. The map preferentially refers to commercial parts, provided by silicon foundries. In addition, there is a significant number of proprietary ICs and ASICs, which were developed by sensor manufactures at their own expense, are exclusively used for the in-house production and are confidential.

## 7.2 Inductive Sensor with Discrete Evaluation Electronics

Continuing with the field of IPSs, covered by the table in Figure 7.2, the evolution paths go back to the 1950s. A historical publication [2] deals with proximity switches developed to detect the approaching or the removing of metallic parts. The paper reviews several procedures based on magnetic or electromagnetic field detections (bridge, resonance methods, etc.). According to the technological level of the 1950s, sensor electronics use vacuum tubes, namely triode valves [3].

Bell Telephone Laboratories (Bell Labs) did make a public announcement of the transistor at a June 30, 1948 press conference in New York City [4]. The replacement of electronics tubes with solid-state devices occurred relatively early in the 1960s.


DCU
MEM
CLG
FSM
PGC
FIF
Cnt
Adr
S#1
S#2
I/O
PSP
• +VCC, +VDD, -VSS
• SURGE- , ESD-, EFT- protection circuitries
+V_B
GND
I/O-BUS
ISE
FEE
BEE
n
AMP
DET
SCD
LIN
AOS
V/I-OUT
OSC
f_osc
CMP
CLU
SOD
& / ≥
OUT
Plumber /Target
V
°C
LTM
TCP
I_OUT>I_SC
D/OI
SCP

List of acronyms and initialisms:

**ISE:** **Inductive sensing element**
**FEE:** **Front-end electronics**
OSC: Oscillator
LTM: Local trimming
AMP: Amplifier
DET: Detector
TCP: Temperature compensation
**PSP:** **Power supply and protections**
$V_{CC}$: Supply voltage
$V_{DD}$: Supply voltage
$V_{SS}$: Supply voltage

**BEE:** **Back-end electronics**
LIN: Lincarisation
SCD: Signal conditioning
AOS: Analog output stage
CMP: Threshold comparator
CLU: Combinational logic unit
SOD: Switching output driver
SCP: Short circuit protection
D/OI: Display/optical indicator
$I_{OUT}$: Output current
$I_{SC}$: Short-circuit threshold

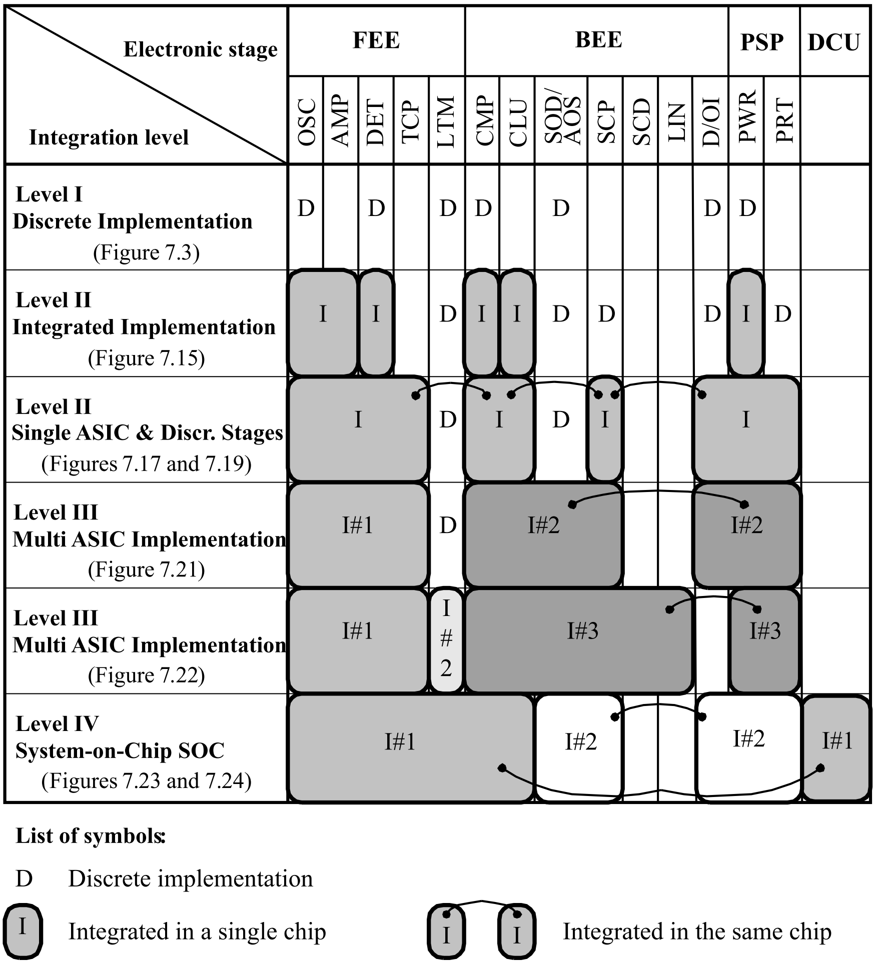
**DCU:** **Digital control unit**
CLG: Clock generator
FIF: Fieldbus interface
FSM: Finite-state machine
MEM: Memory
PGC: Programming and control
Adr: Memory address
Cnt: Memory content
S#1: State #1,
S#2: State #2
I/O: Input/output

**Figure 7.1** Functional block diagram of the generic structure of IS evaluation electronics.

The production of inductive sensors with solid-state devices (discrete transistors and diodes) beside of R-, C-, and L-components started in the 1970s. Such sensors are no longer in serial production; they are now part of history.

Nevertheless, in keeping with the encompassing character of this book and to highlight the simplicity and features of discrete EE schematics, Figure 7.3(a) shows a typical circuitry with bipolar transistors. Similar versions were implemented on a large scale until the 1990s.

The ISE (see Figure 7.1) consisting of the coil $L_1$ and capacitor $C_1$ is supplied by the transistor $T_1$ forming together a Hartley oscillator [5]. The desired positive feedback is provided by secondary coil $L_2$ and the capacitor $C_2$. This classical oscillator will be described in detail in Chapter 8. Its output signal (collector of $T_1$) goes



| Electronic stage / Integration level | FEE | | | | | BEE | | | | | | | PSP | | DCU |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | OSC | AMP | DET | TCP | LTM | CMP | CLU | SOD/ AOS | SCP | SCD | LIN | D/OI | PWR | PRT | |
| **Level I**<br>**Discrete Implementation**<br>(Figure 7.3) | D | | D | | D | D | | D | | | | D | D | | |
| **Level II**<br>**Integrated Implementation**<br>(Figure 7.15) | I | | I | | D | I | I | D | D | | | D | I | D | |
| **Level II**<br>**Single ASIC & Discr. Stages**<br>(Figures 7.17 and 7.19) | I | | | | D | I | | D | I | | | I | | | |
| **Level III**<br>**Multi ASIC Implementation**<br>(Figure 7.21) | I#1 | | | | D | I#2 | | | | | | I#2 | | | |
| **Level III**<br>**Multi ASIC Implementation**<br>(Figure 7.22) | I#1 | | | | I #2 | I#3 | | | | | | | I#3 | | |
| **Level IV**<br>**System-on-Chip SOC**<br>(Figures 7.23 and 7.24) | I#1 | | | | | | | I#2 | | | | I#2 | | | I#1 |

**List of symbols:**

D Discrete implementation

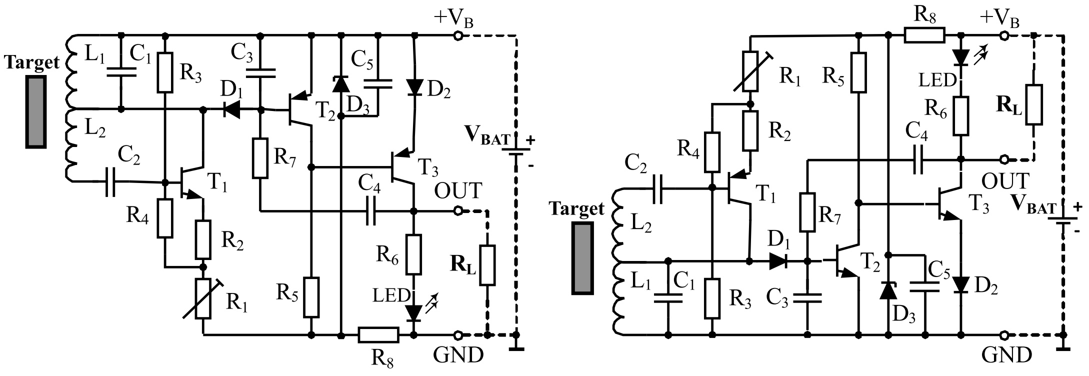
I Integrated in a single chip

I I Integrated in the same chip

Figure 7.2 Technological evolution steps of the evaluation electronics for IPS. PWR = power, PRT = protections. See Figure 7.1 for additional acronyms.

to the easiest peak-detector ($D_1$ and $C_3$). The resulting DC voltage with a sufficient low ripple voltage is led to a Schmitt trigger ($T_2$, $T_3$, and feedback network $C_4$ and $R_7$). At the same time, $T_3$ is the *SOD* of the sensor and can drive an external load $R_L$ connected between the output OUT and ground GND, and thus it provides an HSD operation (see Table 1.2). The dashed representation symbolizes the sensor application: supply source $V_{BAT}$ and sensor load $R_L$ (see Section 1.3.7.1).

The schematic has a minimum number of auxiliary functions. The LTM function is carried out by the adjustable resistor $R_1$ (potentiometer or trimming resistors), which controls the oscillator output signal and finally trims the sensor effective operating distance $s_r$ (see Section 2.1.1.2). The LED fulfills the optical indication. It goes ON if $T_3$ is switched on and the output OUT is active. In order to make possible the operation in the standard supply voltage range, the electronics contain a shunt regulator consisting of the Zener diode $D_3$ in parallel to the bypass capacitor $C_5$ and the series current limiting resistor $R_8$.



(a) (b)

**Figure 7.3** Discrete implementation of evaluation electronics for IPS: (a) HSD version, and (b) LSD version (external circuitry: dashed line = supply and load).

Figure 7.3(b) shows the complementary version for LSD operation, which results in an overturning of the schematic sketch in Figure 7.3(a) with respect to a virtual horizontal line and from the replacement of the npn-transistors by pnp-transistors and vice versa.

## 7.3 Inductive Sensor with Integrated Evaluation Electronics

For a better understanding of the new features and benefits achieved with the implementation of integrated circuits in the inductive sensors, this section starts with two sections (7.3.1 and 7.3.2) that deal with the traditional integration technologies also applied for sensor electronics and summarizes the base knowledge necessary and sufficient for the goal of this book.

The development and fabrication of monolithic ICs and later ASICs with large-scale integration (LSI) for implementation as evaluation electronics (EE) for ISs is an innovation of the 1980s and 1990s. In these decennia, silicon-based bipolar transistors were predominantly used for IC- and ASIC-realizations. They were planarly structured on silicon slices (wafer) with a reasonable diameter between 4 and 6 inches.

The benefits of the integration of the inductive sensor's inner life are specific for this fabrication technology:

- Integration allows substantial improvements of the functionality and makes possible smart, intelligent new sensor features in a comparable housing volume.
- The area for the sensor inner electronics becomes smaller. The components are substantially smaller and the wiring between them needs less area (the majority of connection paths are internally integrated in ICs).
- The sensors are therefore continuously miniaturized and need fewer materials for mechanical parts.

- Performance enhancement is very important regarding sensor reliability. Therefore, the reduced number of electrical connections and conducting traces on the electronics baseboard substantially increases the intrinsic reliability with its criteria MTBF and MTTF (see Section 2.6.1).
- Integration reduces the sensor price, decreasing the material costs as well as the production costs.

### 7.3.1 Overview of Bipolar Integration Technology

The most used material for integrated electronic circuits is silicon (Si), which is the element number 14 in the periodic table. It has 2 + 8 + 4 = 14 electrons on three distinct orbits. The outermost orbit determines the valence of the element, and silicon has consequently a valence of 4. If an atom with a valence of 5 is inserted into silicon, four of its five electrons are tied up with the four electrons of the Si atoms. The fifth electron is unused and free to travel. The doped silicon becomes n-type (n denoting the negative charge of the excess electrons). The more heavily the silicon is doped, the more conductive it becomes.

Similarly, it is possible to introduce atoms with a valence of 3. Now, an electron is missing, forming a positive charge, whose charge, or hole, also moves under the influence of an electric charge. This semiconductor is the p-type.

The way silicon raw material is transformed into final electronic circuits able to be used in inductive sensors involves engineers, metallurgists, chemists and physicists and can be divided into two sections: front end, wherein integration of the electronic circuitry into silicon crystal occurs, and back end, wherein the IC is embedded in inductive sensors, and is done partly at the IC manufacturer and partly at the IC end user (sensor fabrication line).

#### 7.3.1.1 Front-End Processing

Front-end processing is a planar integration process that takes place in specific clean rooms of the IC foundry (since even the smallest particle can ruin the integration) and consists of successive photolithographic steps at specified atmospheric and temperature conditions. Class 1 (one particle of dust in a cubic foot of air) and class 10 clean rooms are typical. Gradually planar layer structures result (Figure 7.4), which contains individual active electronic components (transistors, diodes)

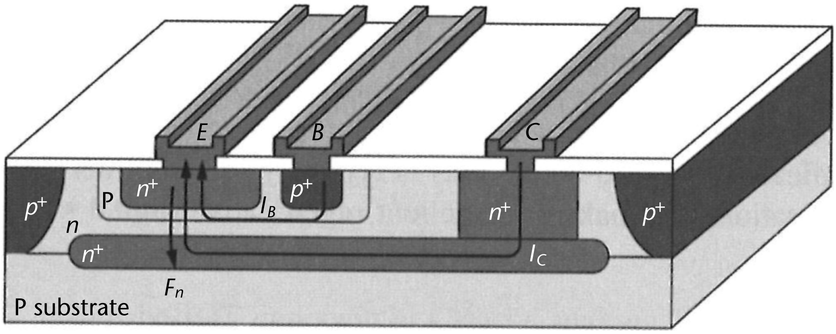


**Figure 7.4** Graphical 3-D sectional view of a monolithic bipolar integration.

but also buried passive components (resistors, capacitors). This section describes these photolithographic steps.

The lowest width of an integrated entity (diffusion or metal conductive strip), called critical dimension, is an important parameter of the integration process. In the 1980s its value was 5 to 10 $\mu$m and in the 1990s it was 0.5 $\mu$m. Due to advancements in semiconductor manufacturing, this process characteristic has been reduced, over all processes, from about 20 $\mu$m in 1960 to about 0.02 $\mu$m in 2012 [6]. A critical dimension of 1 $\mu$m is a competitive value for current bipolar processes (e.g., VIP10 process, manufacturer National Semiconductor).

The intrusion of electronics and computer technology into people's lives and the devices they use daily is growing at an exponential rate, and the well-known Moore's law still applies in the electronic world [6].

The starting material of front-end processing is the raw wafer.

Wafer results slicing (diamond *sawing*) a pure, single-crystal cylinder (called an *ingot*) into thin discs. After lapping and polishing steps, the wafer has the necessary thickness; the common thickness value is 250 $\mu$m. Wafer diameter has continuously increased over time (Figure 7.5) and became possible due to permanent technological progress in semiconductor manufacturing.

The reason for these efforts and massive investments in equipment was to increase the yield of the fabrication batch, which normally contains 20 wafers. The production costs are batch-related and an increase in the number of the functioning ICs obtained significantly lowers the chip price.

To demonstrate yield evolution, Figure 7.6 shows the expected number of good parts (right vertical axis) for an evolutional wafer diameter. The applied chip area is 5 mm$^2$. From the author's experience this is a reasonable die area for IC used in sensors for robust industrial applications.

The electrical characteristics of the wafer are set by doping impurities (dopants). Doping the wafer with p-type acceptor impurities (most commonly boron (B) and gallium (Ga)) to the crystalline silicon results in an increase in the number of majority carriers, namely the holes. Similarly, doping with n-type donor impurities (phosphorous (P), arsenic (As), and antimony (Sb)) leads to a increase in the number of the majority carriers, which are now electrons.

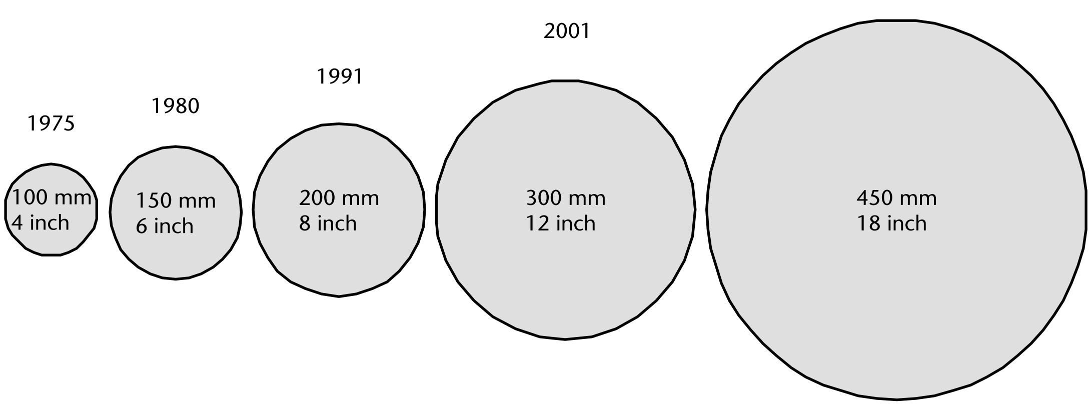


**Figure 7.5** Expansion of the wafer diameter [6].

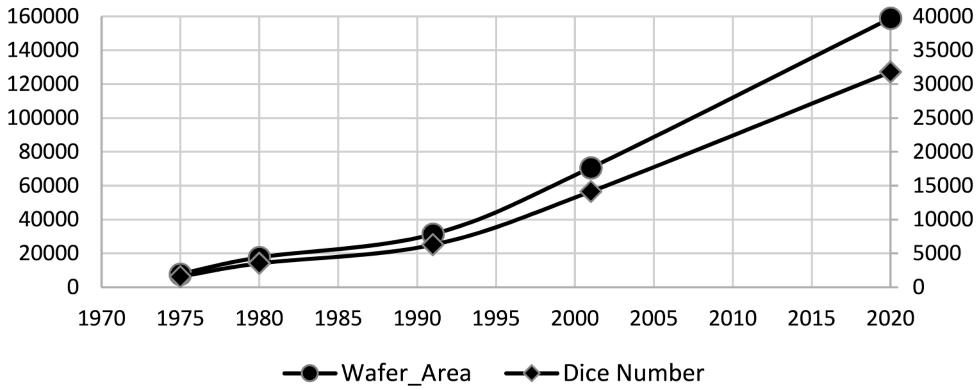


**Figure 7.6** Increasing of the wafer surface area in mm$^2$ (left ordinate) and the IC number (right ordinate).

In equilibrium state, the number of holes or electrons is equal to the concentration of p-type or n-type impurity atoms. The typical impurity concentration of a wafer is $10^{16}$ atoms/cm$^3$.

The outcome of the front-end processing is compact, filigree structures with dimensions (length and width) in the range of a few millimeters (Figure 7.7) and which have a regular geometrical arrangement on the wafer (Figure 7.8).

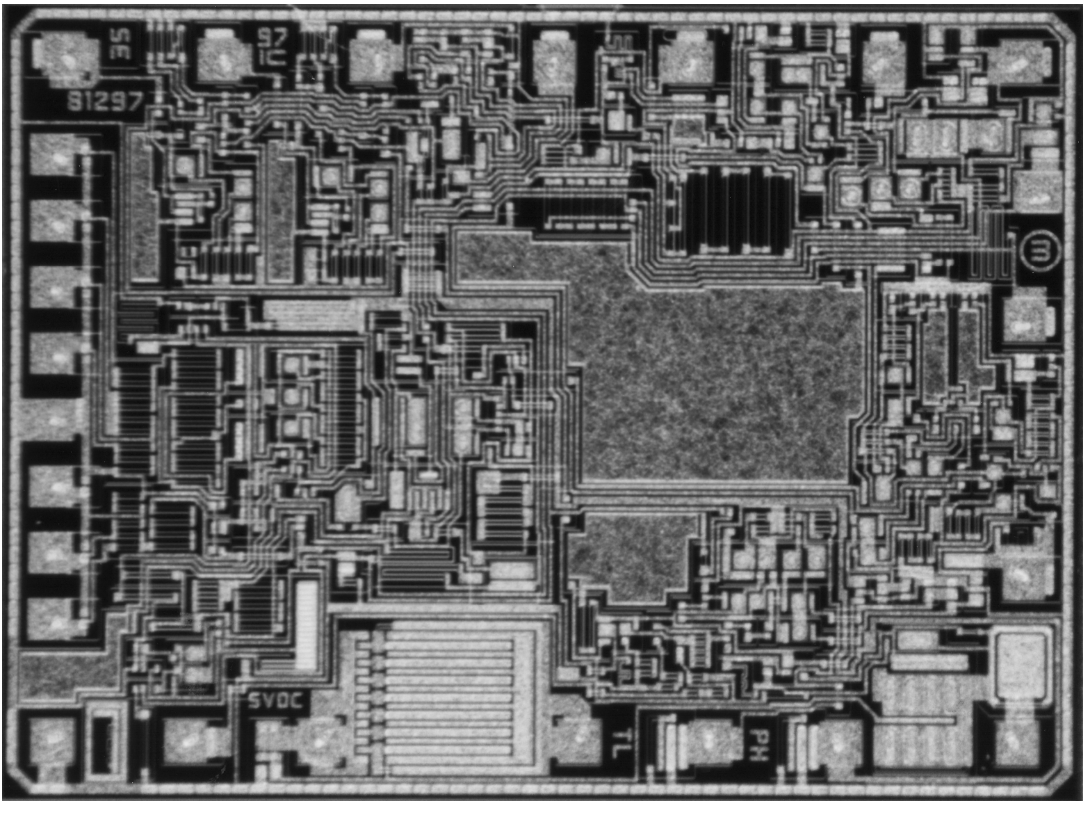


**Figure 7.7** Microscope photography of the active face of a planar-integrated ASIC, fabricated in silicon by means of a bipolar technology.

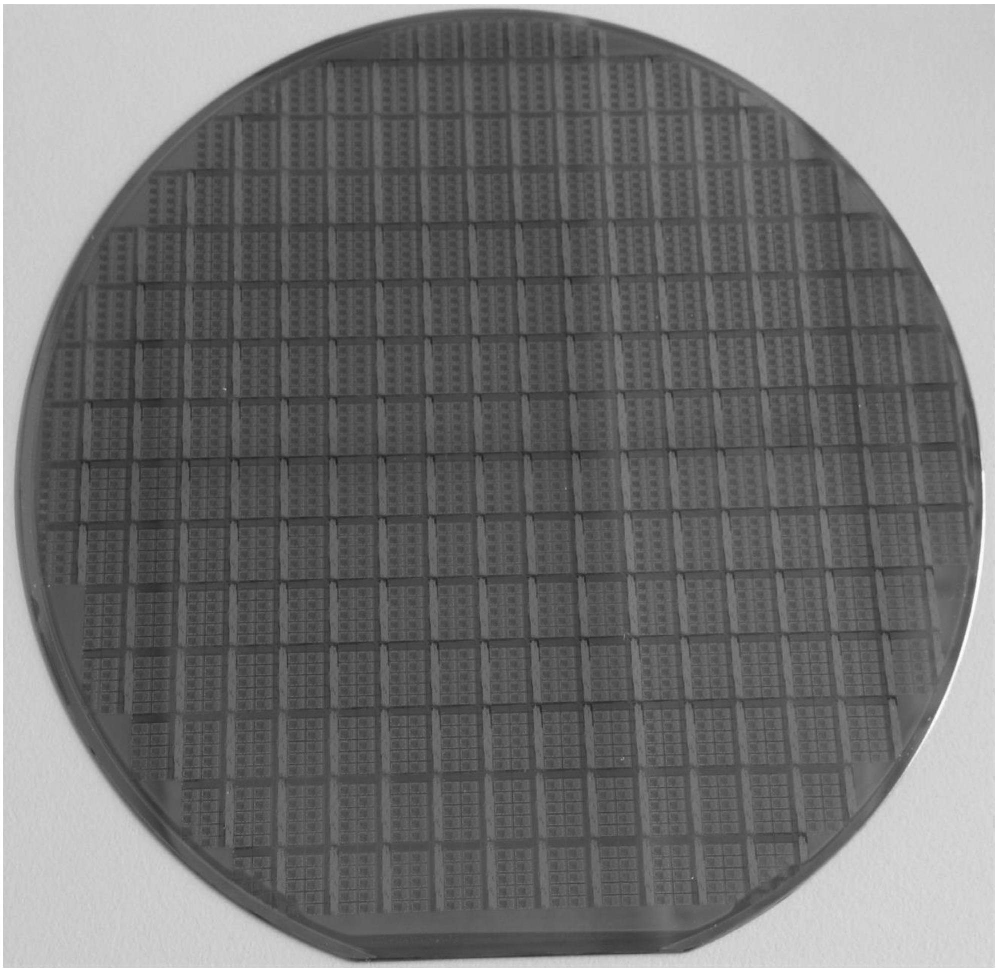


**Figure 7.8** Wafer with integrated entities from the front-end processing (the wafer has one orientation flat edge).

Starting here, the following description outlines the fabrication of a bipolar integrated circuit (for the serial numbers and letter, see Figure 7.9) [7]). The phases are

1. *Thermal oxidation*. This marks the starting of the fabrication. Because the fabrication phases are generally not selective, every phase starts with a window defining operation, which is preceded by an oxide deposition.
   a. The wafer is initially covered by a silicon dioxide ($SiO_2$) layer, which acts as a barrier to the diffusion of impurities. The layer is formed in silicon material, exposing the silicon wafer to oxygen at temperatures of 900°C. Its usual thickness values are 0.2 to 1 $\mu$m.
2. *Buried layer and epitaxial growth*.
   b. The first window for buried layer is provided by performing the following photolithography steps:
      - *Coating*. The photoresist, a light-sensitive protective material, is uniformly applied on the wafer while it is spinning and cured in place (see

Section 5.4). As is generally known, the photoresist becomes insoluble in light-exposed areas.

- *Aligning.* A reticle (glass photomask #1) is positioned on the top of the wafer or projected without touching the wafer. In the next step, light-opaque, dark patterns, which correspond to the necessary windows, will prevent the light from reaching the photoresist;
- *Light exposure.* Ultraviolet light shines through the clear portions of the reticle. The exposed resist polymerizes at exposed areas (outside of tub regions), gets a hard adherence, and remains on the wafer. The wafer pattern is exposed onto the photosensitive resist. Current processes use direct wafer stepping DWS. The mask is quite a bit farther away from the wafer and its magnified image is placed onto the wafer through a series of optics. Because the image is larger than the final pattern, a better resolution can be expressed onto the photoresist.
- *Photoresist developing.* The resist is developed, which means the unexposed resist is washed away and the necessary window becomes visible on the $SiO_2$ layer;
- *Dry etching.* This removes the silicon oxide not protected by resist and opens the window on the silicon using gas instead of chemical etchants, which only attack the oxide and stops at the silicon.
- *Wet etching.* The polymerized resist is stripped to reveal the patterned oxide layer. The batch of wafers is dipped into an acid pool. The method is low cost; however, it does not allow for small geometries.

c. *Solid-state diffusion* is the conventional method to provoke movement, usually at high temperature (typically 1100°C for 1 hour), of impurity atoms from the surface of the silicon sample into the bulk material. The resulting doped silicon behaves electrically as n-type or p-type depending on the impurity. Diffusion is almost always done in a furnace with a flow of gas running over the wafers. The depth of the diffusion varies between 0.1 and 20 $\mu$m. To generate the buried layer a highly-doped ($n^+$)-diffusion process is accomplished through the window before realized.

d. In order to perform the epitaxial (epi) growth, the wafer surface is completely cleared from oxide. Initial planar integration processes used only diffusion steps. To improve the resulting triple-diffused bipolar transistors, epi growth was developed and integrated into the process flow. The process consists of the formation of a single-crystal silicon (epi-layer) on the surface of the silicon sample. Its impurity concentration can independently be controlled and is in opposition to the substrate impurity type. The epi growth of the layer is accomplished by placing the wafer at elevated temperatures in an ambient atmosphere containing chemical combinations based on silicon (e.g., tetrachloride ($SiCl_4$)). A chemical reaction takes place in which elemental silicon is deposited on the entire surface of the wafer. The next layers will be implemented in this crystalline, lightly doped n-type epi-overlayer whose thickness varies between 1 and 20 $\mu$m. This value is decisive for the maximal operation voltage of the integrated parts (e.g., 17 $\mu$m is necessary for a 36V transistor, according to [7]).

   e. To prepare the next phase, the surface of the epi overlayer is recovered with $SiO_2$ (see step a.).

3. *Trench wall processing.* For a better separation between tubs (the cells for transistors, resistors, capacitors, etc.) and, hence, better electrical features of the integrated components like withstand voltage and parasitic capacitances, the diffusion cells in the epi-layer should be isolated by junction walls.
   f. The new windows defining the surface area of the walls are photolithographically opened by means of the mask #2 (see lithography step b. previously described) and a very deep (p+)-diffusion is performed (see diffusion step c.). The concentration of p-type dopants in these areas is higher than the previous concentration of the n-type epi-layer, so that these regions are converted from n-type to p-type. Thus, the tub is electrically separated by junction isolation.
   g. The surface of the epi overlayer is re-covered with silicon dioxide to prepare the next phase.
4. *Base processing.* Using the mask "Base," phase 3. is repeated again to generate the transistor base region in the tub. The process now uses a different p-type of impurities. The base depth measures 1 to 3 $\mu$m.
   h. See step f.
   i. See step g.
5. *Emitter processing.* Using the mask Emitter, the steps h. and i. are repeated at two different coordinates to generate the transistor emitter and collector regions. The impurities are now the n-type. The emitter depth measures 0.5 to 2 $\mu$m.
   j. See step f.
   k. See step g.
6. *Contacts processing.* The structures of the transistor in the tub are now complete; however, the electrode regions of all transistors have to be suitable for interconnection.
   l. Using the mask Contacts and applying photolithography, direct access to these electrode regions are provided.
   m. A total metal deposition is accomplished by *metal sputtering.* In this way, the holes are overflowed with metal layer, which electrically connects all electrodes together. The internal circuit wiring on the device top will be realized in this metal layer during the next phase. The commonly used metal is aluminum. Copper—or at the limit—gold are used only for high-end integration, generally for applications at high frequencies. They involve significantly higher chip prices.
7. *Conductive paths processing.* The conductive tracks of the integrated circuit are made using the mask Alu and the corresponding photolithography.
   n. Using the mask Alu and applying photolithography, the metal layer pattern lies attached on the aluminum deposit. The final metal layer results after the washing of the nonpolymerized resist, after the aluminum etching, and after the stripping of the polymerized resist.
8. *Chip passivation.* In order to have a mechanical and chemical protection of the chip, its surface is covered with a glass or glasslike material, called *passivation.*

o. After a last photolithography with a mask called Pads, the entire chip surface remains open except for the small connection surfaces, called pads, covered by polymerized resist. After the deposition and curing of the passivation material, the resist is definitively stripped. The wafer is in its final state with open pads to be connected to the external chip application or housing.

The described fabrication flow corresponds to the standard traditional bipolar process and requires a minimal number of seven masks and photolithography steps. The IC foundries offer options with additional masks and costs to realize a double metal layer, which considerably simplifies the routing of the conductive tracks, or to get particular features and components as dielectrically isolated tubs with a $SiO_2$ tub border instead of junction isolation (silicon oxide isolation SOI-process option), sinker to access the buried layer, and in this way, to provide a low-resistance path for the NPN-collector, ion implant resistors, and so forth [8].

#### 7.3.1.2 Back-End Processing

Before the chip separation process starts, every chip on the wafer is in turn electrically connected to a test electronics, which belongs to a test card (probe card). The physical connections are made by means of point-shaped carbide metal needles (probing needles). Preferentially they are able to touch the small square chip pads (typical side 50 to 100 $\mu$m) and to provide a good electrical contact without any unwanted scratching (normal needle marks). After the verification of the contacts, the automatic probing procedure is performed. It consists of a series of electrical chip measurements (100 to 500 steps) that test the chip functionality and specifications but also simulate real applications. The bad parts are indicated either by an ink mark directly on the chip surface (inking, obsolete method) or by a symbol in the virtual electronic map (wafer map).

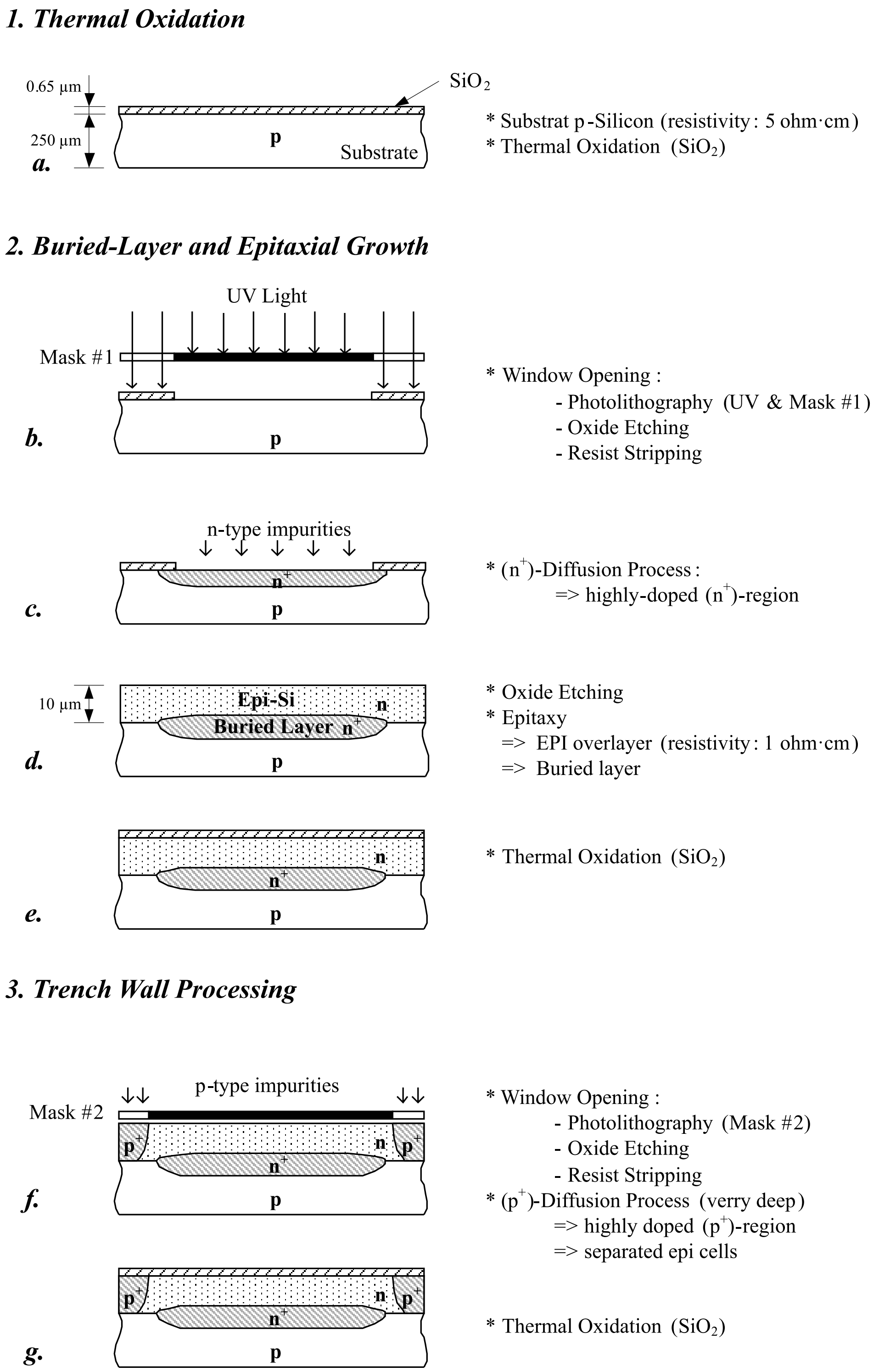
The probing in testing and assembly houses is a time-intensive activity and thus expensive. The probing volume per entity should be a compromise decision to provide the best possible quality and reliability with acceptable costs.

After the probing, the chips on the wafer are physically separated implying mechanical sawing of the wafer by diamond sawing discs or laser beam or by scribing and breaking. The resulting silicon chips are called dice and the singular is die.

The last topic of the back-end is the chip packaging and/or contacting. There are three conventional techniques:

1. *Packaged chip*. This is the most applied procedure, where the a die is fully embedded in a plastic or ceramic housing. There are a large number of specific or standardized packages that contain a cavity in which the chip is placed and a metallic lead frame ending with package pins. Before the die is encapsulated, it is mounted onto the frame, and thin wires connect the chip pads to the frame to create the electrical paths: die pads—frame fingers—housing pins.

   Figure 7.10 illustrates a specific test package (kitpart) to test new chips on testing breadboards. Intentionally, the package cavity is open and the


0.65 µm
SiO$_2$
250 µm
p
Substrate
a.
UV Light
Mask #1
p
b.
n-type impurities
n$^+$
p
c.
10 µm
Epi-Si
n
Buried Layer n$^+$
p
d.
n
n$^+$
p
e.
p-type impurities
Mask #2
p$^+$
n
p$^+$
n$^+$
p
f.
p$^+$
n
p$^+$
n$^+$
p
g.

*1. Thermal Oxidation*

* Substrat p-Silicon (resistivity: 5 ohm·cm)
* Thermal Oxidation (SiO$_2$)

*2. Buried-Layer and Epitaxial Growth*

* Window Opening :
- Photolithography (UV & Mask #1)
- Oxide Etching
- Resist Stripping

* (n$^+$)-Diffusion Process:
=> highly-doped (n$^+$)-region

* Oxide Etching
* Epitaxy
=> EPI overlayer (resistivity: 1 ohm·cm)
=> Buried layer

* Thermal Oxidation (SiO$_2$)

*3. Trench Wall Processing*

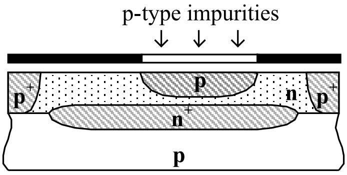
* Window Opening :
- Photolithography (Mask #2)
- Oxide Etching
- Resist Stripping

* (p$^+$)-Diffusion Process (verry deep)
=> highly doped (p$^+$)-region
=> separated epi cells

* Thermal Oxidation (SiO$_2$)

**Figure 7.9** Bipolar standard buried collector SBC fabrication flow sheet (specified typical values).

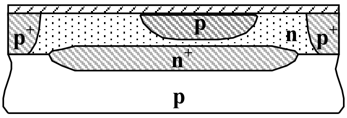
## 4. “Base“ Processing


p-type impurities
Mask #3
p+
p
n
p+
n+
p

*h.*

* Window Opening :
  - Photolithography (Mask #3)
  - Oxide Etching
  - Resist Stripping

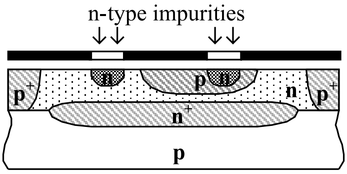
* (p)-Diffusion Process
  => Base-region of the transisitor



*i.*

* Thermal Oxidation ($SiO_2$)

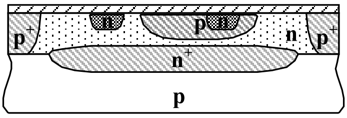
## 5. “Emitter“ Processing


n-type impurities
Mask #4
p+
n
p
n
n
p+
n+
p

*j.*

* Window Opening :
  - Photolithography (Mask #4)
  - Oxide Etching
  - Resist Stripping

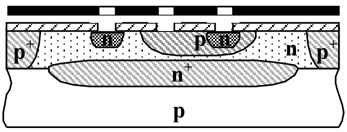
* (n)-Diffusion Process
  => Emitter- and Collector -regions of the transisitor



*k.*

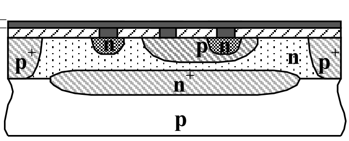
* Thermal Oxidation ($SiO_2$)

## 6. „Contacts“ Processing


Mask #5
p+
n
p
n
n
p+
n+
p

*l.*

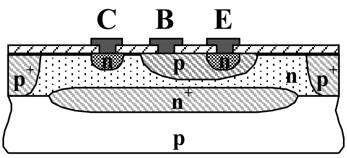
* Window Opening :
  - Photolithography (Mask #5)
  - Oxide Etching
  - Resist Stripping


0.5...1.5 µm
p+
n
p
n
n
p+
n+
p

*m.*

* Metallisation (aluminum sputtering)
  => Total Metal Deposition

## 7. “Conductive Paths” processing.


C
B
E
p+
n
p
n
n
p+
n+
p

*n.*

* Photolithography (Mask #6)

* Aluminum Etching
  => Open Metal Layer

* Resisit Stripping

## 8. Chip Passivation

*o.*

* Inverse Photolithography (Mask #7)

* Glass/Nitrid Deposition => Die Passivation

* Laminar Stripping => Open Pads

**Figure 7.9** *(Continued)*

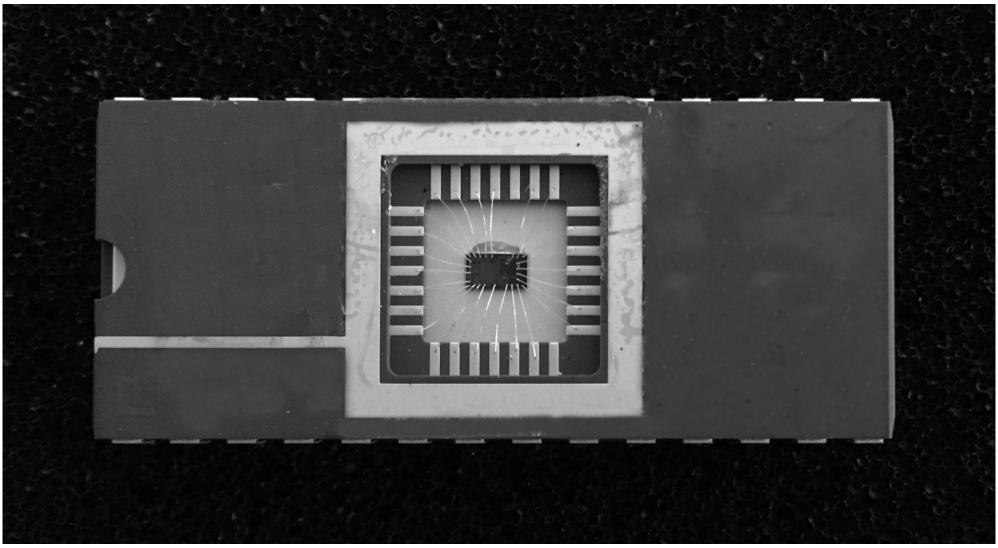


**Figure 7.10** IC chip in open ceramic testing package (visible wire-bonding connections).

photo shows how the chip is assembled. The chip pads are connected to the frame fingers via thin aluminum or gold wire (typical 30-$\mu$m wire gauge). The process is called wire-bonding and the bonds are obtained by a highly accurate, miniaturized welding process.

For encapsulation, molten plastic material is pressed around each die to form its commercial package. The chip is loaded into trim-and-form machines where the lead pins are formed. During encapsulation, product samples are taken out from the normal product flow for environmental and reliability assurance testing. Finally, the packaged ICs are marked and put into antistatic tubes for handling and transportation to the test area for final testing. The final tested devices are available for sale.

The various positions and shapes of the pins as well as the package size and shape depend on the final application. The popular IC packages are suitable for the surface mount technology (SMT). The dual-in-line (DIL) version with pins arranged in two lines like in Figure 7.10 is removed from serial production.

2. *Chip on board* is a novel technology that saves money and allows compact assembly of the sensor electronics. The bare chip is attached and pasted on the sensor carrier PCB with its active face (face with the connecting pads) upstairs (face-up procedure). The electrical links between chip pads and contacting pads on the PC4 are made by wire bonding (Figure 7.11). Finally, the entire zone (chip, wires, and pads on PCB) is covered with a hardened protection compound (glob-top).
3. *Flip chip on board*. Initially developed by the International Business Machines Corporation (IBM) for ceramic carrier electronics boards, the method was adapted and also implemented in the low-cost electronic production segments, which use FR-4 material (see Section 5.4) for carrier boards. The method

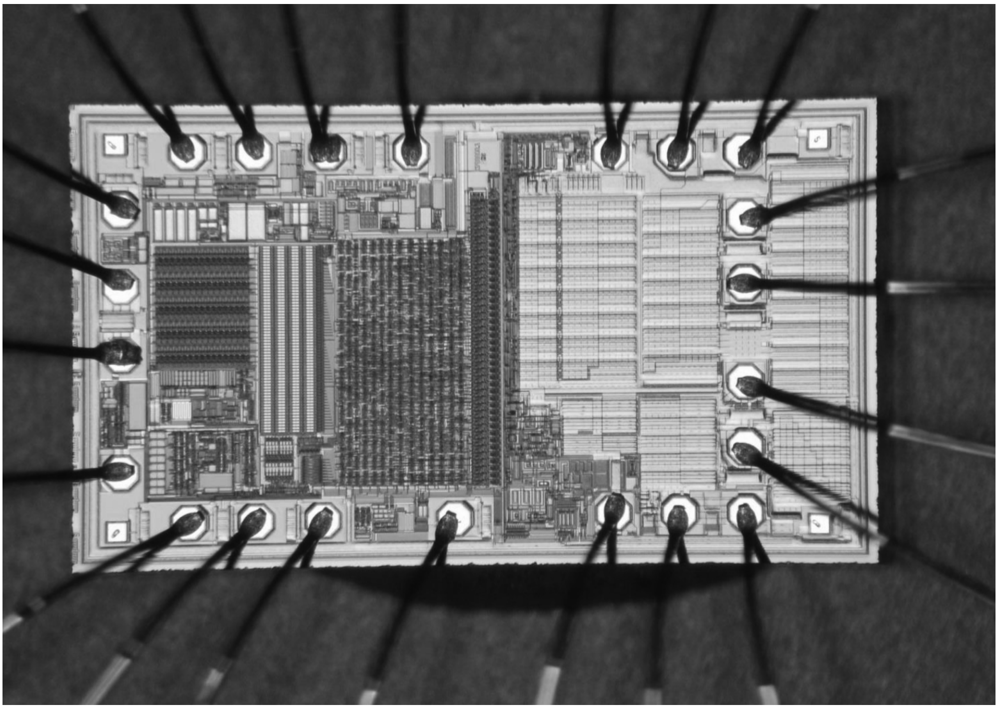


**Figure 7.11** Face-up chip on board with wire bonding connections.

is applied for bare chips with solder bumps [9] as well as for IC in a chip-scaling package [10]. It is a state-of-the-art method that provides extremely miniaturized sensors in parallel with high performance and reliability.

For flip chip applications the tested wafer with probed dice (see the beginning of Section 7.3.1.2) goes to the bumping process. Bumps, which generally have a hemispherical shape and are made from lead-tin alloy (soldering bumps) or gold (gold bumps), will be attached on the chip pads (by a grow-up or electroplating process). After the bumping and a probing after bumping, the wafer is sawn. The attachment of the already bumped chip on the PCB occurs by means of the flip chip method. A pick-and-place machine picks the chip from a chip magazine gadget, rotates the chip, and places it with the active face (bumps face) downward (face-down procedure). The PCB carrying the attached chip as well as the electric parts finally go through a reflow oven. A very stable assembly results (Figure 7.12). Since there are

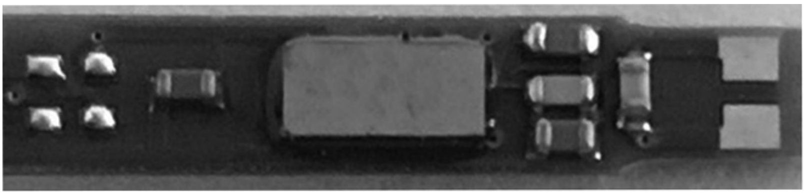


**Figure 7.12** Face-down flip chip on board. Only the gray chip backface (substrate) is visible.

no connection wires and the chip lies on the PCB with its substrate face upward, there is no need for a globe top.

Similar to the chip on board, the assembly of SMD components and the soldering in a standard reflow oven are carried out by the sensor manufacturer and not by the IC supplier.

### 7.3.2 Overview of Complementary Metal-Oxide-Semiconductor Processes: Benchmarking CMOS versus Bipolar Technology

Since the end of the 1990s, the field of integrated circuits for sensor applications the bipolar technology was increasingly replaced by complementary metal-oxide-semiconductor (CMOS) technology. This allowed concurrent engineering for analog and digital circuits on the same chip by means of analog/digital mixed-signal processes. This advancement was an important step for modern sensor electronics, which contain both analog and digital sections (see Figure 7.1).

The CMOS structures are similar to bipolar integrated structures (see Figure 7.13).

Essentially, CMOS processes, which are now the dominant processes worldwide in the semiconductor sector, consist of similar fabrication steps to those in Section 7.3.1.1, with some specifics:

a. Using p-type wafers as a starting material (n-type is also possible), the first step is the formation of a well of opposite impurity-type material, where the complementary device will be formed. The procedure consists of masking and implementation of a donor dopant.
b. A buried layer of metal or polysilicon is deposited and the gate of n-channel devices are defined with a new masking operation.
c. A masking operation is performed so that photoresist covers the p-channel devices and the wafer is etched to an open-source and drain areas of the n-channel devices. Donor n-type impurities are inserted in these areas (diffusion or ion implantation).
d. After a short oxidation, steps b and c are repeated for well devices (p-channel device) using p-type acceptor impurities.

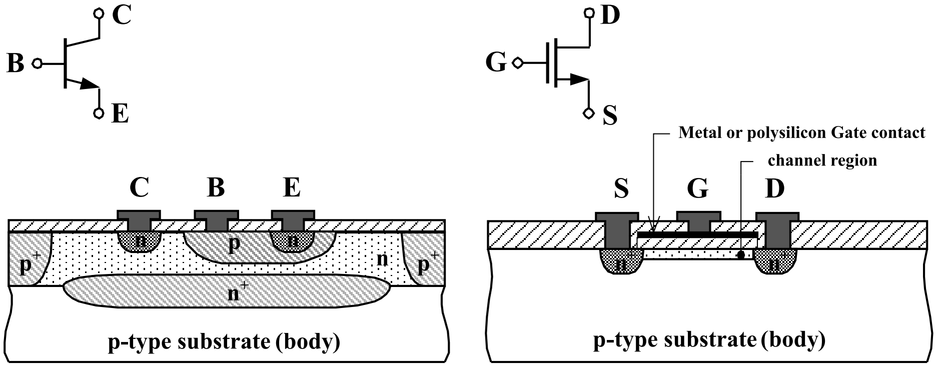


**Figure 7.13** Cross sections of idealized bipolar NPN device versus insulated-gate NMOS device and related symbols for schematic diagrams.

e. A final layer of silicon dioxide is deposited to reduce the parasitic capacitances of the interconnect metallization. The thermal grown of this layer is prohibited to prevent atom redistribution. An alternative to this method is chemical vapor deposition.
f. The contact windows are formed carrying out a masking operation and the metallization is deposited and defined with a second masking step.

MOS technologies have some specific attributes that have to be mentioned:

1. CMOS technologies make possible the integration of NMOS- and PMOS field effect transistors (FETs) on the same chip. FETs can be virtually wattles-switched and the chip current consumption results are very low.
2. For high-accuracy processing, the diffusion is usually replaced by ion implantation. This is a technique for directly inserting impurity atoms into the wafer and shooting the desired dopant ions into the wafer.

   High-energized ions (60 to 80 keV) of the desired impurity species and with a controlled dosage ($10^{14}$ to $10^{16}$ $\mathrm{cm}^{-2}$) are directed at the sample at high velocity. The ions penetrate the wafer surface to a depth up to 1 $\mu$m. The wafer is held at a moderate temperature (e.g., 800°C) for 10 minutes to allow the ions to become mobile.

   The technique is widely applied in MOS technology where small regions with well-controlled amounts of impurity are required. Contrary to diffusion, where the diffusion chamber can handle many wafers, ion implantation can only process a single wafer at a time. At the end, the wafer is heated to make the ions go deeper into the wafer (drive-in step) and to repair the crystal lattice structure damaged by the ion implantation (annealing step).
3. CMOS processes allow to yield digital memories on silicon substrate [6]. They can seize the entire chip area or can occupy a part of this area, besides integrated analog and/or digital circuitries. While the first option has its application domain in the computer world, the second option has more relevance for integrated circuits for industrial sensors. The integrated memory types are in this case:
   - Random access memory (RAM);
   - Read only memory (ROM);
   - Static random-access memory (SRAM);
   - Ferroelectric Random Access Memory (FRAM);
   - Electrically Erasable Programmable Read-Only Memory (EEPROM).
4. CMOS processes can be extended to BiCMOS processes, which are = combined processes making possible CMOS components as well as bipolar transistors. Thereby it is possible to combine the advantage of small structures in the digital area with the superior features offered by the bipolar transistors. Such processes are ideal for mixed-signal applications.
5. CMOS processes support the integration of high-power double-diffused DMOS transistors and combine the advantages of small structures in the CMOS area with the superior switching features offered by DMOS transistors (high currents and voltages, low ON-resistance).

6. Available CMOS processes from different IC foundries span a wide spectrum of complexity and corresponding of selectable photomasks. Figure 7.14 shows several sets of masks available for the processes XC06 [11], which is adequate for integrating sensors evaluation electronics due to its analog/digital/memory-compatibility, good critical dimension (0.6 $\mu$m), and high withstand voltage suitable for industrial applications.

### 7.3.3 Evaluation Electronics of Inductive Sensors with Integrated Circuits

Figure 7.15 shows a possible concrete embodiment of the generic block diagram in Figure 7.1 with commercial, general-purpose ICs. The schematic shows the benefits listed at the beginning of the Section 7.3.

The OSC in Figure 7.1 is realized now with the parallel resonant circuit L and $C_1$ (in fact the inductive sensing element (ISE) in Figure 7.1) and the op amp $OA_1$. The schematic corresponds to the bridge oscillator and will be revisited in Chapter 8. The op amp is preferably a type for single-supply voltage and with rail-to-rail input/output operation. Consequently, there is no need for dual supply and the specific single-supply operation of the sensors for industrial applications is intrinsically fulfilled. A good choice could be MCP6051 (manufacturer Microchip Technology Inc.).

The local trimming (LMT) is carried out by the resistor $R_1$ (similar as for the schematics in Figure 7.3), which has to be adjusted to satisfy the equation:

$$R_1 = R_P \cdot \frac{R_2}{R_3} \tag{7.1}$$

where $R_P$ is the equivalent parallel resistance of the resonant circuit (see (4.38)).

The DET that follows the OSC here is a detector with $OA_2$, which converts the AC oscillator signal (output of $OA_1$) into a DC signal. Its output voltage (emitter of $T_1$) has a DC value equal to the amplitude of the oscillator signal and a small ripple, which can be adjusted to ensure the system stability. The schematic basically corresponds to a peak detector, which consists of a voltage-controlled current source ($OA_2$ and $T_1$) for charging the capacitor $C_2$ and a simple current sink (resistor $R_4$) for discharging of the $C_2$. The $C_2$ is initially loaded up to the peak value of the oscillator signal and periodically—from wave to wave—reloaded to this peak value. By approaching the target to the coil, the oscillator signal goes down and the detected signal has to take part in that process, which is the reason for discharging part $R_4$. This resistor has to be designed to fulfill two conditions: (1) short response time at high speed target approaching ($R_4$ value should be low to accelerate the discharging of $C_2$) and (2) low ripple of the detected voltage (in contrast, $R_4$ value should be higher to reduce the discharging current). The design can be improved replacing $R_4$ by a current sink, which provides a linear discharging from wave to wave.

After detection, the signal path in Figure 7.1 leaves the FEE and enters the BEE.

The CMP is now a regenerative Schmitt trigger [1] with op amp ($OA_3$ and $R_5$ to $R_8$). The symbol of the threshold voltage source $V_{TH}$ stands for a low impedance version (e.g., a voltage divider). This stage converts the DC detector signal into a binary signal offering a sensor switching behavior with hysteresis (see Table 1.2, Section 2.1.1.3, and Figure 2.3). The schematic is analyzed in detail in Chapter 9. The triggering voltages are symmetrically centered on $V_{TH}$ and the hysteresis is a linear function of the limited output voltage levels provided by the two Zener diodes (one biased in a reverse direction and the other in a forward direction). Fully integrated Schmitt triggers can also be used.

The next stage after CMP is the CLU, which basically imposes the sensor switching function (see Sections 1.3.7.1 and 2.1.1.4). In Figure 7.15 this unit has the easiest embodiment; it is a switch (SW) (e.g., a soldering jumper of the electronics carrier board). For the NC switching function (SW position as shown in Figure 7.15), the single inverter gate IG is bypassed. In contrast, for the most-used NO switching function the gate is inserted into the signal paths and it provides an inversion of the switching function. The variety of inverters is very large; a good choice is a single inverter gate from the series SN74xxx04 (manufacturer Texas Instruments Inc.).

The series circuit consisting of the current regulator diode (CRD) and the LED diode forms the stage D/OI in the generic block diagram. The CRD is a less familiar electronic part and needs to be explained. Essentially, it is a junction field effect transistor (JFET), having a series resistor R in its source path (Figure 7.16(a)). Consequently, the transistor volt-ampere forward characteristic presents a horizontal segment (constant $I_F$ current) spread for $V_F$ between 1 and 2V and 100V [12]. The reverse characteristic is not utilized. During the reverse mode, the reverse current IR evolutes along a quite vertical segment as soon as the reverse breakdown voltage $V_R$ achieves a low threshold (see the third quadrant of the graphic in Figure 7.16(b)).

The advantages of the CRD usage are (1) that the LED is supplied with constant current independent of the sensor supply voltage $V_{BAT}$ and therefore its emitting light intensity remains constant, and (2) that the next switching transistor $T_2$ is also driven with constant basis current.

The next stage in the signal path, the SOD, is now realized with a single-housing planar Darlington NPN transistor arrangement ($T_2$ and $T_3$). The medium-power SMD housing includes the bias resistor $R_9$, which speeds up the turn-off action offering a short evacuation path for the majority charge carriers from $T_3$. A breakdown protection diode for $T_3$ is also inside.

The collector of $T_3$ forms the effective switching output of the sensor. However, a positive temperature coefficient thermistor (PTC) is inserted between this collector and the real sensor output OUT. This thermistor is the easiest implementation of the SCP from the Figure 7.1. An exemplifying PTC is the POSISTOR (manufacturer Murata Manufacturing Co.) [13], which increases its resistance value when its temperature rises. This resistance has a negligible value (below 1Ω) at room temperature and increases to above 10 kΩ if the PTC temperature achieves 150°C. This overproportional temperature behavior provides a simple but efficient overcurrent protection. Its limits and up-to-date overcurrent and short-circuit protections are presented in Chapter 10.

The third basic functional unit in Figure 7.1, PSP is represented by the three-terminal positive fixed voltage regulator 3T_PFVR, which supplies the entire electronics with +5V. Recommended parts LM340 or LM7805 [14] are bipolar monolithic integrated circuits (see Section 7.3.1.1) that comply with the standard supply demands for industrial inductive proximity sensors (see Sections 1.3.1 and 2.1.1.5.1). The schematic is fully integrated and need only two ceramic bypass capacitors $C_3$ and

$C_4$. The regulator output has two safe internal protections: (1) a thermal overload protection, and (2) a short-circuit current limitation.

According to the generic functional diagram in Figure 7.1, the stage PSP contains the circuitries for several protection functions. In the present case, there is only one circuitry, the diode array $D_4$ to $D_6$, which performs the sensor surge protection (see Section 2.2.4). Suitable transient voltage suppressor diodes (TVSDs) have a breakdown voltage that is higher than the maximal supply voltage $+V_B$ when they are forward biased and a defined and stable clamping voltage in reversed biasing, even for very high currents (up to the order of amperes).

In fact, for the applied surge pulse across every two sensor connection terminals ($+V_B$ and GND, $+V_B$ and OUT, or OUT and GND) and with both polarities, a series of two TVSDs is present and this limits the pulse voltage on the diodes clamping voltage. This clamping voltage—by a proper design—should be lower than the specific breakdown voltages of the integration processes used to fabricate the ICs.

To recapitulate the operation of the inductive proximity sensor with the electronics represented in Figure 7.15, let us presume a very large distance between the target and the sensing inductor L. The signal of the oscillator ($OA_1$) has the maximal amplitude and the detected voltage ($OA_2$ and $T_1$) has the highest value (higher than the low threshold of the comparator ($OA_3$). Hence, the output of the comparator is LOW. After an inversion in the invertor gate (IG), (programmed NO function), the output of the IG is high and the Darlington transistor $T_2$-$T_3$ as well as the sensor output OUT are turned off. There is of no current through the load $R_L$.

The target approaches the sensor coil and enters into the sensing range. A transition occurs at the moment if the detected voltage reaches the low threshold: the comparator output goes high, the output of IG goes LOW, and the flow of the basis current of the $T_2$ through CRD and LED is activated. As a result, the LED lights up, the Darlington transistor $T_2$–$T_3$ switches on, and the sensor output OUT are turned on. In addition, a voltage step of the internal threshold of comparator happens, from the low threshold to the high threshold. The situation persists if the target furthermore approaches the coil.

If the target moves away from the coil, the switchback does not happen until the detected output voltage reaches the high threshold. In this moment a reverse switching takes place and the turned-off state of the sensor output OUT is restored.

The schematic in Figure 7.15 is symbolized on Level II in the integration levels chart in Figure 7.2.

A remarkable specific of this grade of integration is the possibility to take a look and to analyze the schematic of the integrated circuits down to the transistor level (see for example the inner life of the LM340 in [14]). The manufacturers generally disclose the schematics and their medium complexity allows them to be understood.

An alternative to the implementation of more IC—normally commercial parts—is the use of the analog front-ends (AFE). They are monolithic integrations that contain either independent classical electronic circuits or an already wired schematic using these stages. The first category is for general purpose, has lower prices, and can immediately be used for inductive sensor circuitry. The second category is designed for a specific application and needs to be adapted for inductive sensors. The market is rich in such devices; some are freely available, while others are proprietary to

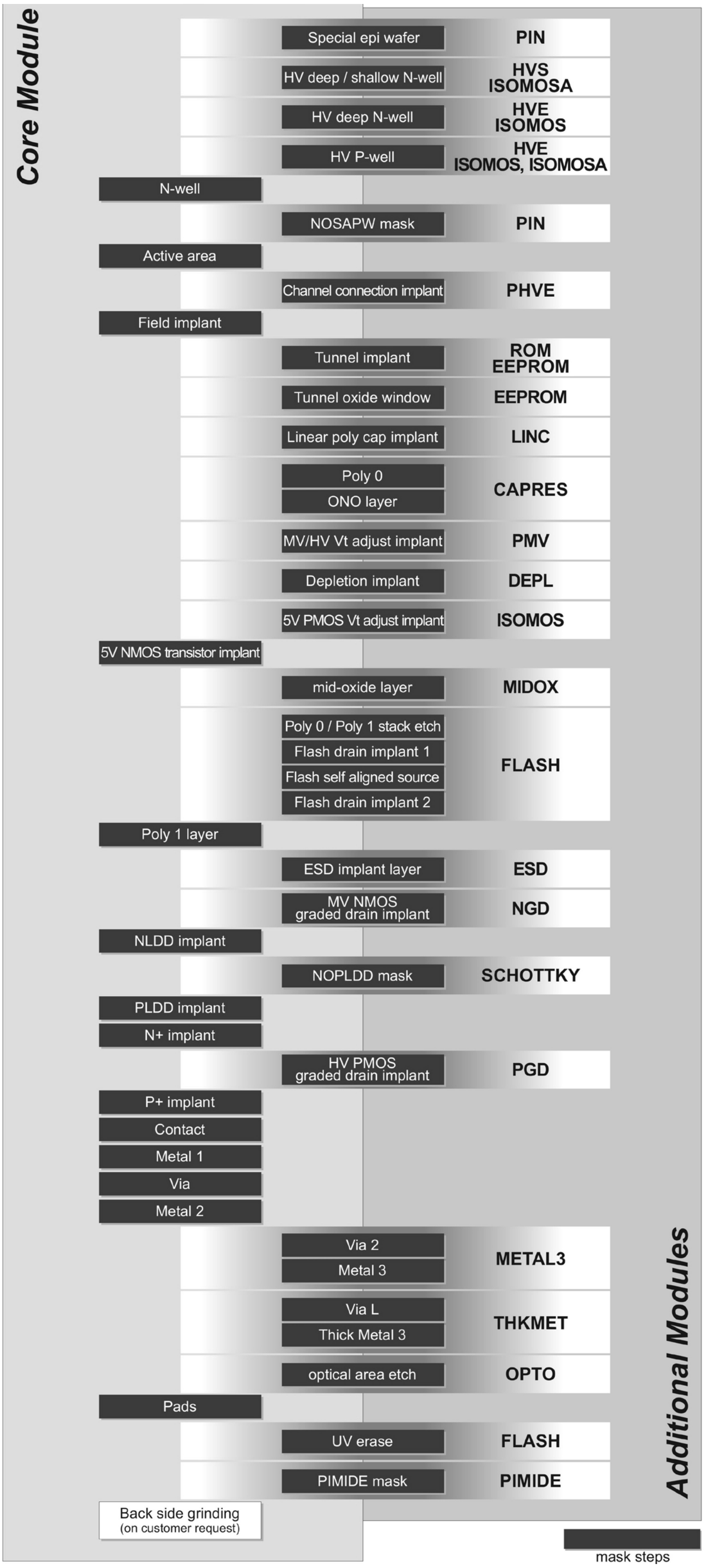

Core Module
Special epi wafer
PIN
HV deep / shallow N-well
HVS ISOMOSA
HV deep N-well
HVE ISOMOS
HV P-well
HVE ISOMOS, ISOMOSA
N-well
NOSAPW mask
PIN
Active area
Channel connection implant
PHVE
Field implant
Tunnel implant
ROM EEPROM
Tunnel oxide window
EEPROM
Linear poly cap implant
LINC
Poly 0
ONO layer
CAPRES
MV/HV Vt adjust implant
PMV
Depletion implant
DEPL
5V PMOS Vt adjust implant
ISOMOS
5V NMOS transistor implant
mid-oxide layer
MIDOX
Poly 0 / Poly 1 stack etch
Flash drain implant 1
Flash self aligned source
Flash drain implant 2
FLASH
Poly 1 layer
ESD implant layer
ESD
MV NMOS graded drain implant
NGD
NLDD implant
NOPLDD mask
SCHOTTKY
PLDD implant
N+ implant
HV PMOS graded drain implant
PGD
P+ implant
Contact
Metal 1
Via
Metal 2
Via 2
Metal 3
METAL3
Via L
Thick Metal 3
THKMET
optical area etch
OPTO
Pads
UV erase
FLASH
PIMIDE mask
PIMIDE
Back side grinding (on customer request)
Additional Modules
mask steps

**Figure 7.14** XC06_masks and main process flow [11].

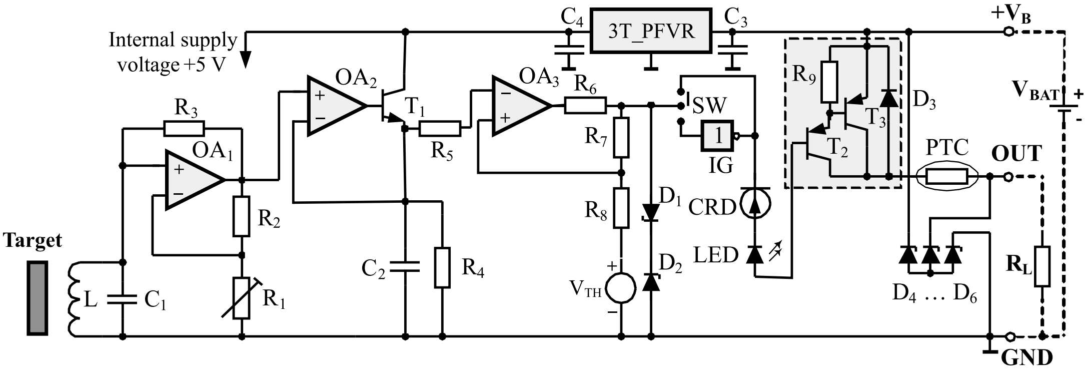


**Figure 7.15** Schematic diagram of the sensor evaluation electronics of a proximity-inductive sensor (version with integrated circuits).

the specific IC foundries, IC design houses, or research institutes and need to be negotiated. The following list provides a few examples: TDC1101 (manufacturer Texas Instruments, www.ti.com), Sentinel (manufacturer ProTek/USA, www.protekdevices.com), EM7180 and EM8500 (manufacturer EM-Microelectronic Marin/ Switzerland, www.emmicroelectronic.com), and HMT7742 (manufacturer HMT microelectronic/Switzerland, www.hmt.ch).

## 7.4 ASIC Implementations in the Evaluation Electronics for Inductive Sensors

It is self-evident that the possibility to apply a monolithic integration technology for EE moreover increases sensor performance and reliability and at the same time decreases the sensor size and expenses.

The major benefits of these high-integration-level ASIC components can be summarized by the following:

- High functionality and flexibility as well as high packaging density;
- Small dimensions and applicability for miniaturized sensor sizes;
- Interoperability with the automation of the sensor assembly lines;
- High sensor feasibility and reliability, since the majority of connection lines are already tested chip-internal lines (metal layer) and not strip lines on PCB;
- Convenient chip price, even considering the nonrecurrent expenses NRE for ASIC circuit design, layout drafting, prototyping, testing, and so forth.

This section contains two parts, which present more implementations of ASICs in the EE of inductive proximity sensors for binary or analog versions.

### 7.4.1 Single-ASIC Implementations: The Classical Device TCA505

The implementations can be assumed as a first milestone: the sensitive part, which is the front-end electronics, is taken over into a monolithic ASIC, while the back-end electronics are furthermore discrete, as in Figure 7.15.

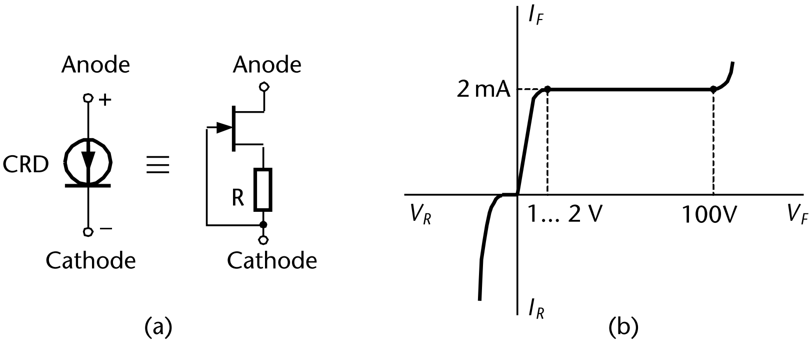


**Figure 7.16** (a) The symbol and the equivalent schematic of a current regulator diode CRD. (b) The volt-ampere characteristic of the part CR200 [12].

An adequate example for this level is offered by the ASIC TCA505BG (manufacturer Infineon Technologies, Germany) [15]. The history of TCA505 begins at the end of the 1970s with the development and fabrication of the TCA105. The concept was continuously improved by Siemens AG, Semiconductor Group and every three to four years a new series part came on the market: TCA205, 305, 505, 605, and 705 [16]. The most successful part, TCA505, is applied in a total volume of many millions and is still available today after 30 years. Such volumes, long-time availability, and use are a specific characteristic of successful ASICs for industrial applications.

TCA505 forms a complete front-end electronics and additionally provides some back-end specific functions (power supply, short-circuit protection, etc.). It is compatible with several back-ends (discrete or integrated). The reason for its renown: it provides just enough implementation (without special features), which fulfills standard demands.

In contrast to the ICs, which can be analyzed on the transistor level, ASICs can be better analyzed using a functional block diagram and schematics of some blocks. Using the full schematic of an ASIC is laborious and senseless.

The block diagram of TCA505 (with manufacturer notation and terminology) and a minimal number of peripheral components are shown in Figure 7.17 [15]. The signal channel: oscillator–demodulator–threshold switch is, in this case, a series signal processing chain and is slightly different from the generic parallel version in Figure 7.1. The reason is the oscillator of TCA505, which is involved in the excitation of the resonant LC circuit contained in the ISE as well as in the capture of the output signal $v_{LC}$ of this circuit.

The change in amplitude of $v_{LC}$ as a result of the target influence is transmitted to a threshold switch by means of a demodulator and finally triggers the ASIC outputs. A distance resistor ($\geq 300\Omega$) between the pin $R_{DI}$ and ground sets the current in the oscillator circuit. The greater the value of the resistor, the smaller the current feed from the oscillator into the resonant circuit LC and therefore the greater the switching distance. The greater the quality factor of the resonant circuit, the greater the value of the resistor $R_{DI}$ necessary for setting a certain switching distance $s_r$ (see Section 2.1.1.2).

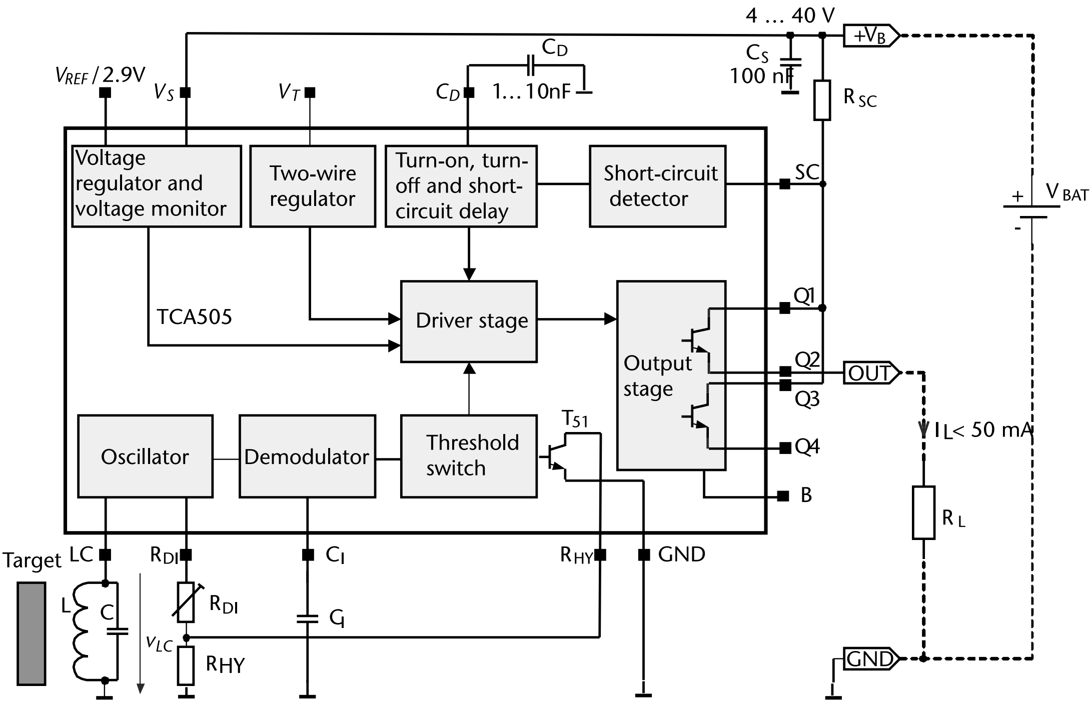

4 … 40 V
+VB
VREF/2.9V
VS
VT
CD
CD
1…10nF
CS
100 nF
RSC
SC
+
-
VBAT
Voltage regulator and voltage monitor
Two-wire regulator
Turn-on, turn-off and short-circuit delay
Short-circuit detector
TCA505
Driver stage
Output stage
Q1
Q2
Q3
Q4
B
OUT
IL< 50 mA
T51
Oscillator
Demodulator
Threshold switch
RL
Target
LC
RDI
CI
RHY
GND
L
C
vLC
RDI
RHY
CI
GND

**Figure 7.17** Block diagram of TCA505 with manufacturer notation and terminology [17, 18].

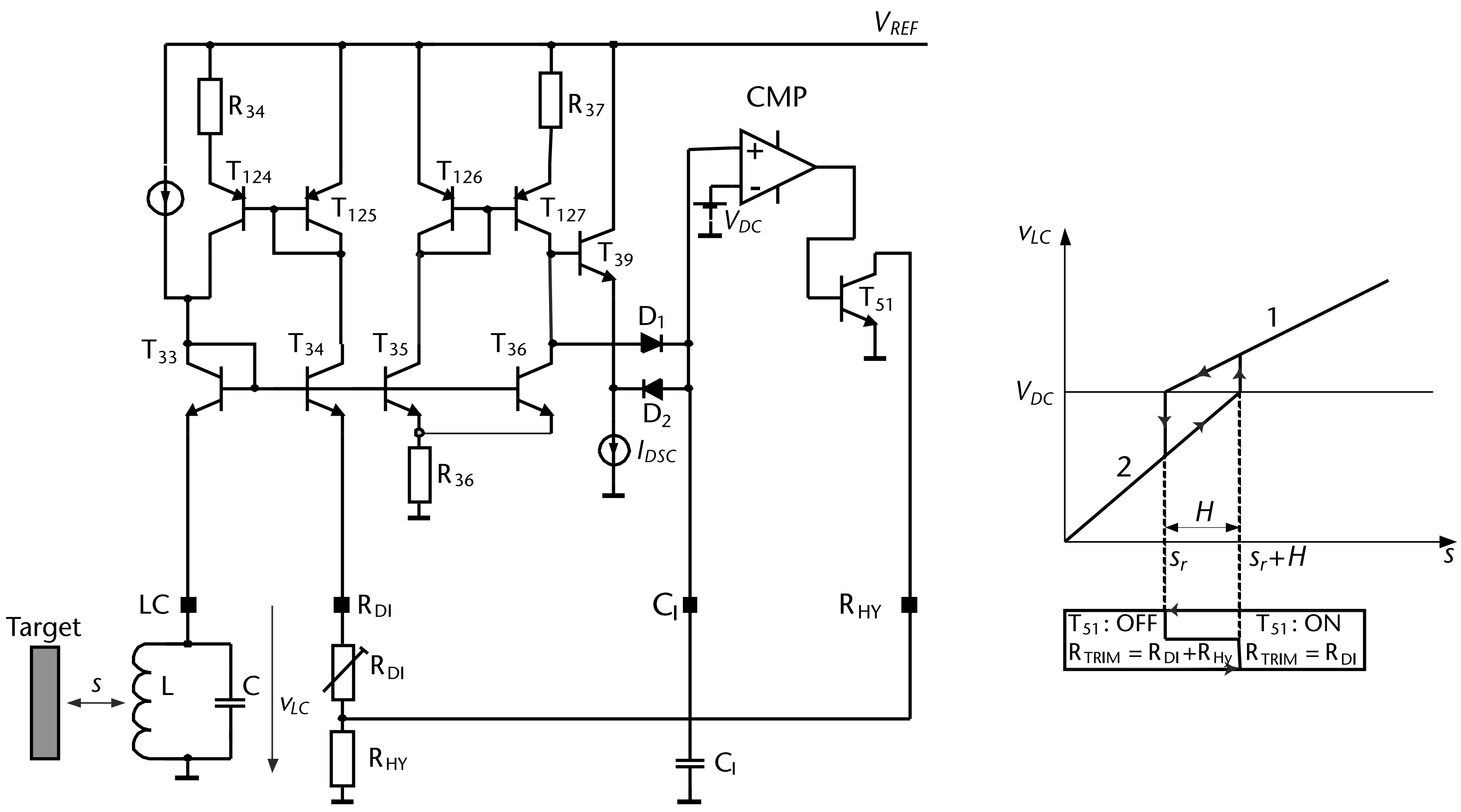
The schematic in Figure 7.18(a) is a basic representation of the electronics of the above-described single channel. The transistors $T_{33}$, $T_{34}$, $T_{125}$, and $T_{124}$ draw a *ring-oscillator* structure. The signal $v_{LC}$ actuates the voltage-controlled current source $T_{34}$, whose output collector current is mirrored in the current mirror $T_{125}$ and $T_{124}$ and is reinjected without phase rotation in the LC (positive feedback). The circuit is analyzed in detail in Chapter 8. Anticipating this analysis, the condition to have steady-state oscillations is:

$$R_{DI} \approx \frac{R_P}{2} \tag{7.2}$$

where $R_P$ is the equivalent parallel resistance of the resonant circuit (see (4.38). To satisfy the (7.2) the trimming resistor $R_{DI}$ has to be adjusted (see the LTM stage in Figure 7.1).

The primary oscillator signal available on the collector of $T_{33}$ is fed to the demodulator (transistors $T_{35}$ to $T_{39}$), which has a similar operating mode like the peak detector with $OA_3$ in Figure 7.15. The capacitor $C_I$ will initially be loaded through the diode $D_1$ up to the peak value of the oscillator signal $v_{LC}$ and periodically—from wave to wave—reloaded to this peak value. To provide a good reaction speed on the target movement, the capacitor $C_I$ is discharged from peak to peak through the diode $D_2$, and in this case, through an internal current sink $I_{DSC}$.

The manufacturer recommends a series RC circuit ($R_I$, $C_I$) if high noise immunity is to be achieved.


$V_{REF}$
R$_{34}$
R$_{37}$
CMP
+
-
T$_{124}$
T$_{125}$
T$_{126}$
T$_{127}$
T$_{39}$
$V_{DC}$
T$_{33}$
T$_{34}$
T$_{35}$
T$_{36}$
D$_1$
D$_2$
T$_{51}$
R$_{36}$
$I_{DSC}$
Target
LC
R$_{DI}$
C$_I$
R$_{HY}$
s
L
C
$v_{LC}$
R$_{DI}$
R$_{HY}$
C$_I$
$v_{LC}$
$V_{DC}$
1
2
H
s
$s_r$
$s_r$+H
T$_{51}$ : OFF
R$_{TRIM}$ = R$_{DI}$+R$_{HY}$
T$_{51}$ : ON
R$_{TRIM}$ = R$_{DI}$

(a) (b)

**Figure 7.18** (a) Schematic circuit diagram of the oscillator, demodulator, and threshold switch [17]. (b) The composite dependence $v_{LC}$ versus $s$ and the state diagram.

The detected signal is led to the stage called threshold switch, which consists of a CMP followed by the open-collector switching transistor $T_{51}$. The commutation process and the hysteresis generating are illustrated in Figure 7.18(b). For high distances *s* between target and the coil L, the peak value of the voltage $v_{LC}$ exceeds the comparator reference $V_{DC}$, $T_{51}$ is switched on and short-circuits the hysteresis resistor $R_{HY}$. According to lower trimming resistance $R_{TRIM} = R_{DI}$, the oscillator transfer function (i.e., $v_{LC}$ versus *s*), is schematically represented by its segment 1. The target approaching causes $v_{LC}$ to go down until the value $V_{DC}$ is reached.

That happens at the already adjusted effective operating distance $s_r$. Now, an active commutation takes place. Because the voltage $v_{LC}$ is lower than the comparator reference $V_{DC}$, the comparator turns off the transistor $T_{51}$, the resistor $R_{HY}$ is released, the trimming resistance increases to $R_{TRIM} = R_{DI} + R_{HY}$ and the operation point steeply jumps on segment 2 of the oscillator characteristic.

A removal of the target cases a new commutation if $v_{LC}$ becomes equal to $V_{DC}$ again. That happens at a distance greater than $s_r$. The difference corresponds to hysteresis H. In summary, the external resistor $R_{DI}$ serves to adjust the sensor switching distance within wide limits and the second resistor $R_{HY}$ serves to set the required fixed hysteresis.

Besides these basic functions (oscillator, demodulator, and threshold switch), the bipolar monolithic ASIC TCA 505 B includes a number of useful extra functions that enable high-grade inductive proximity switches to be designed for an attractive price/performance ratio and with space savings inside.

The driver and output stages in Figure 7.17 correspond to the circuits CLU and SOD of the back-end functional unit BEE in Figure 7.1.

The output stage has a TCA505-specific structure and consists of two antiphase output stages: Q1–Q2 (for a NO switching function, see Table 1.2) as well as Q3–Q4 (NC switching function) for maximum of 50 mA. The pins Q1 and Q3 or Q2 and Q4 can be connected in parallel as required.

The function of the outputs is ensured when the emitter voltage of the output transistors (pins Q2, Q4) is between 0V and the voltage on the additional pin B. The pin B originally serves to limit the base voltage of the internal output-stage transistors. If B is not connected, the operating range of Q2 and Q4 extends to about $V_S$–2V ($V_S$ is the supply voltage). If B is connected to $V_{REF}$ for example, it is possible to set a constant output current that is independent of the supply voltage by means of an external resistor $R_Q$ across Q2 (or Q4) and ground (Figure 7.19(a)).

The internal output transistors are driven in a floating state, thus providing the user with optimal flexibility for evaluation of the output signals. It is therefore possible to use the internal output transistors either (1) as an emitter follower, (2) as an open collector, (3) as a current source, or (4) in push-pull operation. The external circuitry with dashed lines in Figure 7.17 shows an HSD utilization (see Table 1.2) for load currents below 50 mA.

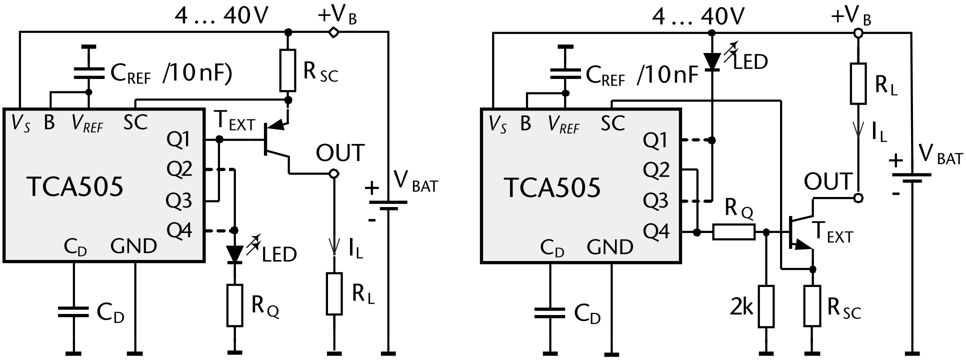
When the pin B is connected to $V_{REF}$, Q2 and Q4 can be used between 0V and $V_{REF}$ and any constant current up to 50 mA can be set on the outputs by means of resistors on Q2 or Q4. This operation mode allows to extend the output current range using an external switching transistor and additionally to connect an LED, which does not have a specific stage in TCA505. Figure 7.19.(a) shows an HSD version in which TCA505 controls an external power pnp-transistor $T_{EXT}$. The base current is adjusted by the resistor $R_Q$ and also supplies the LED. The LSD topology in Figure 7.19.(b) uses an external npn-transistor and has a similar operation.

The stages short-circuit detector and short-circuit delay perform the overload and short-circuit protection of the ASIC output(s). The current sensing is indirectly made by measuring of the voltage drop between $V_S$ and the sensing pin SC or between SC and ground (threshold ≥0.3V). This voltage is generated by means of a sense resistor $R_{SC}$, which is inserted in the output current path as follows:

- For applications with a load current $I_L \leq 50$ mA, the integrated output transistors can be used (see Figure 7.17). The resistor $R_{SC}$ is connected between the supply line $+V_B$ and the collector Q1 (for the HSD version) and the voltage across this resistor is sensed by the input SC. The opposite circuitry with $R_{SC}$ connected between GND and the emitter Q2 (for the LSD version) is also possible.
- For applications with higher load currents, which uses an external transistor $T_{EXT}$ controlled by the TCA505 (see Figure 7.19), the sense resistor $R_{SC}$ is removed in the emitter circuitry of this transistor.

In contrast to the schematic in Figure 7.15, which has the simple short-circuit protection (current limitation with PTC thermistor) having the disadvantage of a long recovery time if the short circuit is turned off, TCA505 has a performant protection circuitry. This accomplishes a pulsing low-energy procedure that does not produce thermal stress during the short circuit; it can protect the sensor for an infinite long time and immediately returns to the sensor normal operation, if the short circuit disappears or is removed. The procedure essentially corresponds to an astable multivibrator.

When the input SC detects an overload or a short circuit, the internal switching transistors Q1–Q2 and Q3–Q4 are immediately turned off. After an adjusted recovery time $t_P$, the transistors are turned on for a very short $t_{SC}$ time in order to check if the ASIC mishandling persists. This time $t_{SC}$ is short enough to prevent



(a) (b)

**Figure 7.19** TCA505 applications with external output transistor $T_{EXT}$, LED indicator, short-circuit protection, and selectable NO/NC switching function (see alternating selection dashed lines): (a) HSD version, and (b) LSD version.

the destruction of the internal and/or external switching transistors. This is also a short useful delay because brief glitches on the outputs or the charging of sensor line capacitances should not trigger the short-circuit protection. If the mishandling does persist, the transistors are returned off. The outputs Q1 to Q4 are already internally protected against overcurrent during the time $t_{SC}$. The current will not exceed 250 mA. The multivibrator is continuously running until the mishandling ends.

The value of the sense resistor $R_{SC}$ depends on the sensor specified maximum load current $I_{L\max}$:

$$R_{SC} = \frac{0.3\text{ V}}{I_{L\max}} \tag{7.3}$$

The short-circuit times are adjustable by means of the delay capacitor $C_D$ according to the following formulas:

$$t_{SC} = 2.5\ \frac{\mu\text{s}}{\text{nF}} \cdot C_D\ [\text{nF}] \text{ and } t_P = 0.5\ \frac{\text{ms}}{\text{nF}} \cdot C_D\ [\text{nF}] \tag{7.4}$$

that denotes a turn-off to turn-on ratio of:

$$t_P = 200 \cdot t_{SC} \tag{7.5}$$

TCA505 provides defined turn-on delay (TonD) and turn-off delay (ToffD) capability under normal operation as well. For a graphical explanation of these times, see Figure 1.7. The delays are also adjustable with the capacitor $C_D$:

$$TonD = 0.65\ \frac{\text{ms}}{\text{nF}} \cdot C_D\ [\text{nF}] \text{ and } ToffD = 25\ \frac{\mu\text{s}}{\text{nF}} \cdot C_D\ [\text{nF}] \tag{7.6}$$

The capacitor $C_D$ has usual values between 1 and 10 nF.

A switching regulator is incorporated to supply the ASIC when it is used as a two-wire inductive proximity switch (see Table 1.3), and this is activated when the pin $V_T$ is connected to $V_S$.

The circuit also has a line regulator generating a stabilized voltage of approximately 2.9V that is brought out on pin $V_{REF}$. The supply voltage range in normal operation is between 4 and 40V. A capacitor $C_{REF}$ can be connected between $V_{REF}$ and ground to improve the noise immunity of the overall circuit function.

If pin $V_{REF}$ is connected to $V_S$, the circuit is operating between 3.1 and 4.5V. In this case, however, $V_{REF}$ is no longer internally stabilized (i.e., the analog IC functions depend on the operating voltage).

The schematics in Figure 7.19 can be extended inserting a surge protection network with suitable transient voltage suppressor diodes TVSD as shown in Figure 7.15.

TCA505 is specifically designed for inductive four-wire switches with antivalent outputs (see Section 1.3.7.1). The extension is immediate. The outputs Q1 and Q3 will be separated from each other and the output circuitry with external transistor $T_{EXT}$ (see Figure 7.19.(a)) will be duplicated for both outputs.

At the end, the TCA505 features can be summarized as follows [15]:

- Low current consumption of less than 0.8 mA typically;
- LC-oscillator working frequency ≤3 MHz;
- Integrated output stage for up to 50 mA output current;
- Temperature range −40 to +110°C;
- High switching frequencies up to 5 kHz;
- Various external drivers are supported: bipolar and MOS transistors, as well as thyristors or triacs;
- Short-circuit and overload protection of output stages and external parts;
- Good temperature coefficient (±2.5% over entire temperature range);
- High noise immunity;
- Suitable for two-wire AC proximity switches;
- Improved temperature drift, noise immunity, and the switching frequency compared to earlier ICs for inductive proximity switches;
- Compact die having a chip area of about 4 mm$^2$ despite high-voltage bipolar technology (Figure 7.20);
- Availability in 16 pin-SMD package (PG-DSO) but also as die for wire-bonding assembling (see Section 7.3.1.2).

### 7.4.2 Multi-ASIC Versions

Circuit combinations between integrated parts and discrete stages lead the way to multi-ASIC versions, which contain two or more ASICs and a minimum number of passive peripheral components.

This makes possible further miniaturization of the inductive sensors evaluation electronics in parallel to the enhancement of the functionality and complexity without additional space requirements and/or costs.

A topology, for example, is suggested in reference [19]. The proposed schematic uses two commercial ASICs from the STMicroelectronics [20]: the monolithic proximity detector TDA0161 and the intelligent power switch TDE1708DFT. Figure 7.21 shows the block diagrams of both ASICs and the interconnections between them, as well as the passive components to accomplish an inductive sensor. This sensor can directly be connected to an industrial application; for example, a high-side driver (HSD) topology (see Table 1.2).

The inner workings of the TDA0161 is similar to the TCA505 (see Section 7.4.1) and, after the previous presentation, a detailed explanation relative to the operating mode is not needed. However, the types of the oscillator and output stage are different. The oscillator is now a Colpitts circuit (see Chapter 8), which requires a resonant circuit with three components and works at frequencies up to 10 MHz. For this reason, the inductive sensing element contains an inductor L and two capacitors $C_1$ and $C_2$ and has three connections to the ASIC: a detector hot point (DET-HP), detector (DET-E) and filtering (manufacturer notation and terminology).

The adjustment of the oscillator signal amplitude and hence of the switching distance $s_r$ is made with the resistor $R_1$. After detection in DET, the output signal of the Schmitt trigger CMP controls the output stage, which is a voltage-controlled current source. The oscillation stops when the target is closed to the L and sustains when the target is far away from the sensor. Independent of the supply voltage $V_{CC}$ (4 to 35V), the output current $I_{OUT}$ is correspondingly high (≈10 mA) or low (<1 mA) according to the presence or absence of the target, respectively.

This monolithic switching device TDE1708DFT is especially designed for various proximity detectors and contains a power switching transistor (PST) (open emitter, open collector) with overload and short-circuit protection, thermal shutdown (for chip temperature >180°C), and protection against kickback glitches, which appear when an inductive load is switched off. A protection signal is internally generated to block the input signal INP and to prevent activation of the output switch, as long as the above-mentioned abnormal conditions are detected. The power-on transition, as well as the chip over temperature and the output overcurrent concur to the generation of such signal. A minimum delay of 25 $\mu$s (typical) is added to the trailing edge of such signal to ensure that a stable normal situation is present when the signal disappears. The delay can be further increased connecting the capacitor $C_6$ between the pin DELAY and GND (adjustment coefficient 1 ms/nF).

The PST can be supplied with high voltages up to 50V and can handle load currents up to 0.5A. Finally, the device can drive resistive or inductive loads in both directions: high-side driver (HSD) (load to GND) or low-side driver (LSD) (load to +supply line).

TDE1708DFT contains a +5V voltage regulator (5V-REG) with a current capability up to 6 mA, which can supply external front-ends. A current driver of a status indicator, LED-DRV, is also available and internally activated if the transistor PST is switched on. It can source or sink LED current (3 mA typ.) according to the chosen output configuration.

Returning to the specific HSD application in Figure 7.21, the generated voltage drop over the resistor $R_2$ connected at the output OUT of TCA0161, having the levels ≤0.33V and 3.3V, is the input signal of the TDE1708DFT. The next Schmitt trigger STG (300-mV fixed hysteresis) in the signal path correspondingly controls the switching transistor PST if the AND gate is enabled by the comparator CMP1. That happens under normal conditions, when the capacitor $C_6$ is charged on about 5V (over the 4V threshold of CMP1).

The capacitor $C_6$ will immediately be discharged through the transistor $T_6$, and PST will be turned off, if (1) the thermal shutdown detects a chip temperature higher than 180°C or if (2) the second collector of PST senses overcurrent or short circuit in the output circuit. The comparator COM2 switches over and turns on the $T_6$.

The integrated LED driver supplies the external LED to indicate the output state. The thresholds of the output comparator CMP3 and CMP4 are set at about 1.5V. For instance, in the HSD case, when the voltage on the output HS differs from GND less than 1.5V, the output is considered in OFF state and the LED driver is disabled. If the pin HS differs from GND more than 3V, then the output is sensed ON and the driver will force the current on the LED.

The power supply bypass capacitors $C_4$ and $C_5$ are not a sine qua non condition. Nevertheless, they improve the sensor stability.

Beside popular devices as TCA505 and TDA0161, worthy of being mentioned are the up-to-date ASICs manufactured by Texas Instruments Incorporated, like LDC0851, LDC1612, or LMP91300 [21], which can similarly be used single-handed or followed by a discrete output stage or by an integrated output driver.

All the previous schematics have a common weak point resulting from the sensor trimming element. This is an adjustable resistor or a potentiometer, which can be handled only if the sensor housing—metallic in general—is open. Closing the housing can result in sensor feature deviations that cannot be compensated for.

The solution is to perform the sensor trimming at the end of the fabrication if the sensor is finally assembled (closed sensor housing). This procedure is implemented in the schematic shown in Figure 7.22, which represents the evaluation electronics of an analog inductive sensor. Besides the tandem arrangement consisting of the analog front-end AFE and the back-end ASIC, the circuitry contains a third part, the trimming *ASIC*.

The AFE is a customer proprietary ASIC (owner Balluff GmbH & Inc., www.balluff.com ) and the representation in Figure 7.22 initiates a transition from the range of commercial ASICs to the confidential area of customer proprietary ASICs.

To signalize this fact, the AFE is represented by its chip photo [22]. The black circles are soldering bumps already attached on the chip pads (see Section 7.3.1.2). The inner workings are less well known; however, its functional structure is very similar to the generic schematic in Figure 7.1 (signal processing chain: OSC, AMP, DET, SCD, LIN, and TCP) [23].

In contrast, the back-end ASIC is a commercial part AM401 with a large application area for industrial proximity sensors [24]. The modular device consists of the following four autarkic parts with individual configuration options:

- A robust voltage regulator/reference, which is suitable for industrial applications and can be supplied in a wide range: $V_{CC} = 6 \ldots 35$V (quiescent current < 1.5 mA). The regulated output voltage *VREF* can be set to 5V if the input *VSET* is not connected (like in Figure 7.22) or to 10V if the input *VSET* is connected to GND. Due to the pnp-transistor in the line regulation, the stage needs a voltage stabilizer load capacitance $C_{REF} = 2.2\ \mu$F, typically.
- An op amp *OP1* for input signals referenced to GND, whose gain is adjustable and which can act as a general purpose analog output stage (AOS in Figure 7.1). Due to a particular design, its output voltage *VOUT* ranges from 0V to ($V_{CC}$–5V) and can thus fulfill the requirements: 0 to 5V or 0 to 10V (see Section 1.3.7.2). In Figure 7.22 the single-ended input voltage of OP1 is the output signal of the AFE, which varies in a low range and has to be amplified in OP1 to cover the standard output range 0 to 10V. Hence the gain of the noninverting OP1 is adjusted to:

$$G_{OP1} = \frac{VOUT}{VINOP} = 1 + \frac{R_1}{R_2} = 4 \tag{7.7}$$

OP1 can internally generate output currents up to 10 mA, can drive low loads down to $R_L = 2\ \mathrm{k}\Omega$, and has an output overload protection by current

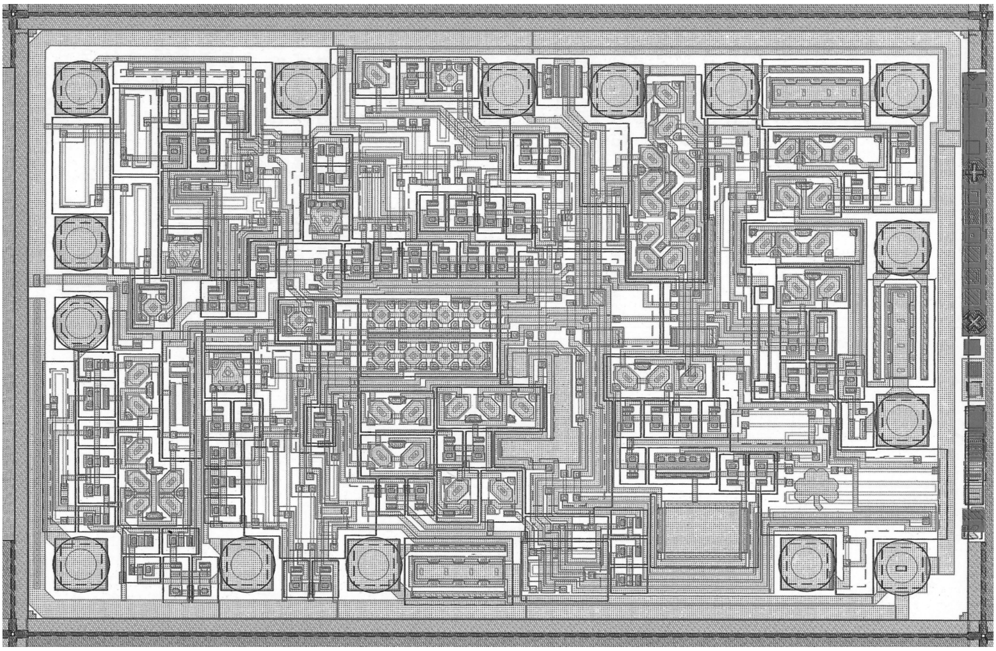


**Figure 7.20** The layout plot of TCA505. Chip area is equal to 4 mm$^2$ [17].

limitation. The terminals $V_{CC}$, $VOUT$, and GND are protected against reverse polarity (see Chapter 11).

- The high-precision instrumentation amplifier (IA) with an internal gain of $G_{IA} = 5$ and the ability to externally set a reference potential (input ZA). The IA acts as an input stage for differential voltage signals (±400 mV) as, for example, the output voltage of an AC deflection bridge (see Figure 3.16).

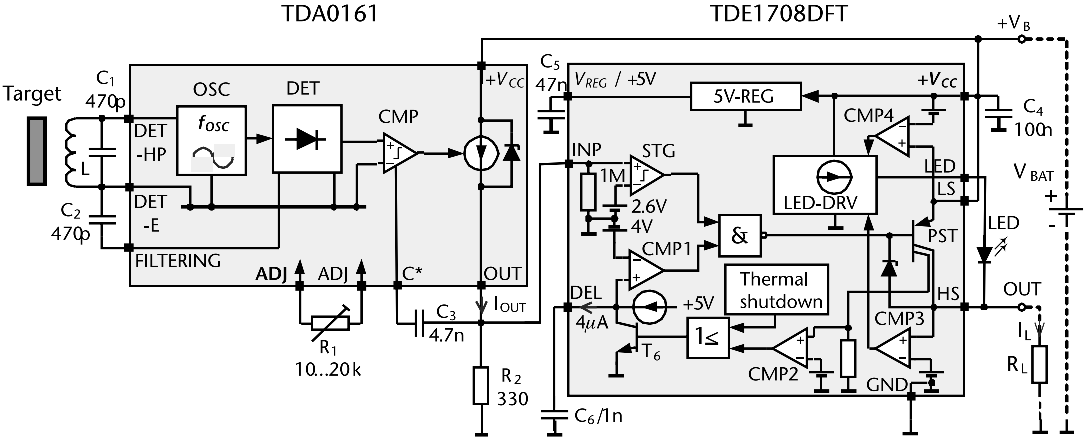

TDA0161
TDE1708DFT
+V_B
Target
C₁
470p
OSC
DET
f_osc
DET
-HP
CMP
+V_CC
C₅
47n
V_REG / +5V
5V-REG
+V_CC
C₄
100n
L
C₂
470p
DET
-E
FILTERING
ADJ
ADJ
C*
OUT
I_OUT
C₃
4.7n
R₁
10...20k
R₂
330
C₆/1n
INP
STG
1M
2.6V
4V
CMP1
&
LED-DRV
CMP4
LED
LS
PST
LED
V BAT
Thermal
shutdown
DEL
4μA
+5V
T₆
1≤
CMP2
HS
CMP3
GND
OUT
I_L
R_L

**Figure 7.21** Inductive sensors evaluation electronics with two ASICs: TDA0161 and TDE1708DFT (block diagrams with manufacturer notation and terminology [20]), interconnections and passive components.

- The second noninverting op amp *OP2* if biased by a bandgap voltage reference $V_{BG} = 1.27$ V, typ. [7] and can be used as a current or voltage source to supply external additional circuits. This output value can be set across a wide range using one or two external resistors, respectively, and has a good stability due to the bandgap consistency.

The stages IA and OP2 are not used in Figure 7.22 and are correspondingly disabled. The ASIC will be reloaded and described in detail in Chapter 10.

A main topic of interest in Figure 7.22 is the third functional unit, the trimming ASIC, which is symbolized in Figure 7.1 by the local trimming (LTM). The concept, embodiment, and programming method are subject of a U.S. patent [25] and the embodiment ASIC is over 20 years in serial fabrication and in use with error-free results.

The presentation of this ASIC is suitable for using the "four-times I" criteria (introduction, implementation, interface, and integration) as following:

1. *Introduction* of the theme. The main component of the trimming ASIC is the resistor $R_X$, which controls the amplitude of the oscillations $v_{\mathrm{LC}}$ of the oscillator in AFE. The device is permanently $\mathrm{V}_{DD}$- supplied from the AFE and has two additional connections to that: its output $\mathrm{R_X}$ and its input called stop trimming ST.
2. *Implementation* of the patent embodiment. Though the $\mathrm{R_X}$ is represented in Figure 7.22 as a simple resistor, it is actually a resistive network (series and/or parallel) whose value is controlled by the electronic switches of the entity switches-memory-counter SW-MM-CT. The increase or decrease of $\mathrm{R_X}$ is a result of a countdown sequence or countup sequence in the counter CT. The volatile content of CT is first saved in the nonvolatile memory MM (e.g., an EEPROM memory), whose bits control the switches in SW. This functionality is managed by the central processing unit (CPU) as soon as a trimming operation is initiated.
3. The *interface* of the trimming ASIC is integrated and its programming input (PGIN) is directly connected to the inductive sensor terminals $+\mathrm{V_B}$ and GND (see Figure 7.22). For a new trimming, the final assembled compact sensor is temporary connected to an external programming unit. According to the Balluff patent [25], this unit generates a programming command that physically overlays the DC supply voltage $+\mathrm{V_B}$. The communication protocol is also customer proprietary. The logical level of this command consists of an 8-bit word whose content codes the different operations (start countdown, start countup, etc.). The bit time duration, modulation method, and carrier signal parameters are suitable for a valid signal transmission from the terminal $+\mathrm{V_B}$ to the input PGIN. The internal transmission gate (TG) precedes the signal and leads it to the CPU. This decodes the received command and accordingly controls the instruction execution in CT-MM-SW. As a result, a start-stop teach-in procedure will be launched.

   Let us suppose the distance *s* between the target and the coil L is mechanically adjusted on the specified high limit $LR_{\max}$ of the linearity range LR (see Figure 1.3) and the received instruction initiates a countup sequence

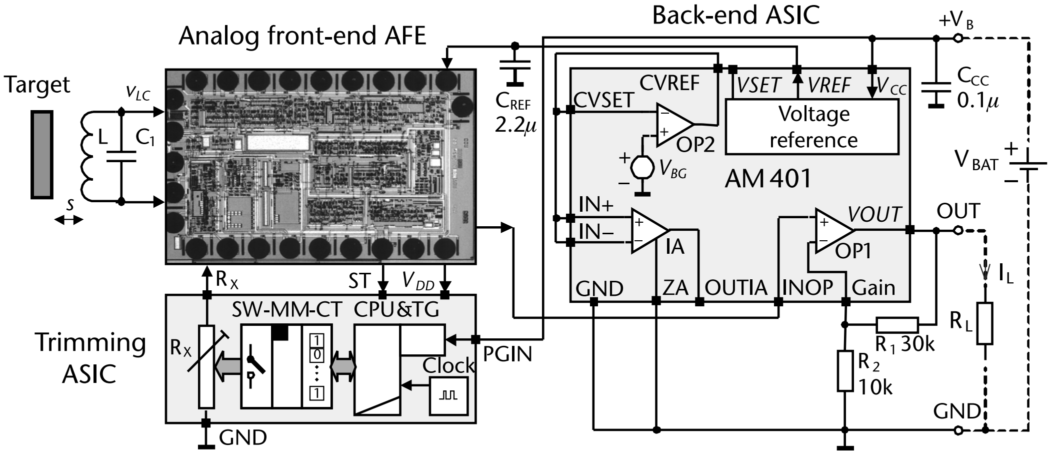


**Figure 7.22** Inductive sensors evaluation electronics with three ASICs: customer proprietary AFE and trimming ASIC as well as a commercial back-end ASIC AM401 [24], interconnections and passive components.

in the counter CT starting with zero CT content. This content goes up and the value of $R_X$ goes down. This decrease of $R_X$ provokes in AFE a step-by-step increase of $v_{LC}$. When this voltage reaches the threshold of an end-of-range comparator in AFE, this stage generates a stop signal realized in a commutation of the signal ST. Hence, the countup in the trimming ASIC is immediately stopped and the momentary value is stored from CT in MM. Therefore, the trimming of the inductive sensor is finished and the programming unit, which only initiated the procedure, can be removed.

4. The *integration* of the trimming ASIC is made using an analog/digital low-power CMOS process with the following features: p-type wafers, n-well, polygate, double metal, number of masks = 12 (see Section 7.3.2). The specific low supply voltage (<5.5V) is not critical because the ASIC is supplied with low voltage from the AFE. The technology is characterized by a relative high critical dimension of 2 μm (see Section 7.3.1.1). The reason for this choice was to ensure a great resilience against the harsh industrial working conditions (voltages, electromagnetic perturbations, etc.). This decision generated a larger chip area (3.7 $\text{mm}^2$) and therefore a higher chip price. However, its correctness is confirmed by a positive field experience accumulated during 20 years of use. In order to allow repeated programming (in the fabrication line, for final tests or fine tuning in the application, etc.), the second strategic decision was to implement an EEPROM memory instead of one-time programming (OTP) circuits (for more details see Chapter 12).

### 7.4.3 Systems on Chip SOC

The high-resolution, high-speed device LDC1101 (manufacturer Texas Instruments Inc.) is really an up-to-date integrated system which performs inductive sensing nearly without passive peripheral components and which offers a broad band of highlights (operating as well as adjusting features). According to its specification

[26], the device covers wide value ranges of resonant circuit parameters: L and C, Q-factor, and working frequency:

- Inductance: $L \in$ [1 $\mu$H to 500 $\mu$H];
- Capacitance: $C \in$ [100 pF to 10 nF];
- Q-factor: $Q_L \in$ [10 to 40];
- Frequency: $f \in$ [500 kHz to 10 MHz];
- Operating temperature: $T \in$ [−40°C to +125°C].

The system can be used for analog and/or binary implementations in inductive proximity sensors. Therewith, in our opinion, the device is able to fulfill all the demands of sensors for industrial applications.

### 7.4.3.1 LDC1101 System Architecture and Operation, Evaluation, and Power Modes

The functional block diagram in Figure 7.23 is a top-level representation with Texas Instruments terminology of the complex LDC1101 architecture [26].

The sensor driver supplies the resonant circuit LC with AC current to maintain steady-state oscillations. It is able to drive currents between 4.7 $\mu$A and 600 $\mu$A in the frequency range and for Q values specified above and simultaneously senses the changes of the LC-circuit parameters. Using the main parameter of a parallel resonant LC circuit, namely the equivalent parallel resistance $R_P$ ((see the definition formula (4.38)), the range which can be covered by LDC1101 is:

$$R_p \in \left[1.25\ \text{k}\Omega \text{ to } 90\ \text{k}\Omega\right] \tag{7.8}$$

The typical value of the oscillating signal $v_{LC}$ (see (4.42)) is 1.2V peak-to-peak and does not exceed 2.3V. LDC1101 reports error condition if the oscillation stops.

The preprocessed signal is transmitted to two independent measuring subsystems, which offer the following two evaluation modes of the magnitudes of the resonant circuit LC and finally provide the evaluation results in digital form:

a. *RP + L measurement* has a double functionality. It simultaneously performs an evaluation and computation of the equivalent parallel resistance $R_P$ as well as of the inductance $L$.

   The subsystem accomplishes this task by regulating the amplitude of the signal $v_{LC}$ in a closed-loop configuration at a constant level while monitoring the energy dissipated by the resonant circuit. By measuring the amount of power injected into the resonant circuit, the device determines the value of $R_P$, which it returns as a digital value with 16-bit resolution.

   In addition, the subsystem can also measure the oscillations frequency $f_0$ by counting the number of cycles of the reference frequency (clock signal (CKL)) and comparing both frequencies. Based on this digital measured oscillations frequency $f_0$ and on the well-known capacitor value C, the stage computes the inductance L (see (4.33)) and delivers it with a 16-bit resolution.

The subsystem has three programmable features:

1. The predicted variation range of $R_P$ as a result of the target actuation, to increase the resolution (see Section 1.3.4) by limitation of this range;
2. The time constants necessary for sensor oscillation;
3. The sample rate for the RP + L conversion. Its typical value is 180,000 samples per second.

b. *LHR measurement (high resolution)* is the second subsystem, which carries out only a high-resolution determination of the inductance L of the attached resonant circuit. The result resolution can be selected: either 24 bit or 16 bit. The resolution is a function of the above-mentioned sample rate and depends on the magnitude of the reference frequency (CKL). The stage can provide a constant conversion time interval; the longer this time, the higher the accuracy.

The subsystems are autarkic and can convert simultaneously but at different sample intervals. The completion of the first conversion is asynchronous to the completion of the second conversion.

The threshold comparison stage can compare the measured $R_P$ with already programmed thresholds (Schmitt trigger functionality [1]). It can provide a switching output signal (SWS) with hysteresis, reported as high/low level on the output line SDO. This function is provided only for RP + L conversion results.

The supply voltage is specified for $V_{DD}$ = 1.8 ... 3.3V and the typical current consumption is strongly dependent on the power mode. A bypass capacitor $C_{LDO}$ = 15 nF must be connected from the pin CLDO to GND.

The system has the following three power modes [26]:

1. *Active mode* with/without connected resonant circuit (typically 3.2-mA or 1.9-mA quiescent current, respectively).
2. Sleep low-power mode (typically 135-$\mu$A quiescent current), in which the LDC1101 retains register settings and can quickly reenter active mode for conversion. The serial peripheral interface (SPI) is functional so that reads and writes can be performed.
3. Shutdown very-low-power mode is a significantly reduced power mode (typically 1.4-$\mu$A quiescent current), in which the device configuration is not retained.

Configuring the LDC1101 must be done while the device is in sleep mode.

The registers and logic unit contains 32 registers (each 8-bit register length) plus the hard-coded logic and is clocked by the signal CKL available at the input CLKIN. The requirements regarding the CKL signal are [26]

- Frequency range: 1 to 16 MHz. For optimum $L$ measurement performance, the highest frequency is recommended.
- Clean and low jitter signal.
- Amplitude swing within the range of $V_{DD}$ and GND.
- Duty cycle between 40% and 60%.
- Rise and fall times below 5 ns.

The registers can be classified into the following groups:

- Configuration registers (hexa-addresses: 0 × 01 to 0 × 05, 0 × 0B, and 0 × 0C);
- $R_P$-threshold setting registers for binary switching applications (hexa-addresses: 0 × 06 to 0 × 0A);
- $L$-threshold setting registers for binary inductance evaluation (hexa-addresses: 0 × 16 to 0 × 19);
- $R_P$ and $L$ conversion results data outputs registers (hexa-addresses: 0 × 21 … 0 × 24 and 0 × 38 to 0 × 3A);
- Operation status reporting registers (e.g., 0 × 0A, 0 × 20);
- Miscellaneous registers (offsets, ID values, etc.).

The bidirectional communication of the LDC1101 with the microcontroller level MCU occurs over an SPI [27]. Originally defined by the former company Motorola, SPI is a simple, low-cost, four-wire, master-slave interface intended for CPU to peripheral communication and control. The meanings of the SPI signals in Figure 7.23 are

- CSB: chip select (active LOW).
- SCLK: clock input.
- SDI: slave data input. It is identical to the master signal MOSI (master out/slave in) (see also Chapter 14).
- SDO: slave data output. It is identical to the master signal MISO (master in/slave out). If CSB is high, SDO goes into a high-impedance state and the line can be driven by the above-defined switching signal SWS.

The packaged device LDC1101 in a VSON housing with a body size of 3 × 3 mm has a GND pin but also a metallic die attached pad DAP on its reverse side (Figure 7.23). This pad is used to improve the thermal performance and is internally connected to the GND. It can be connected externally to the GND or can be left open.

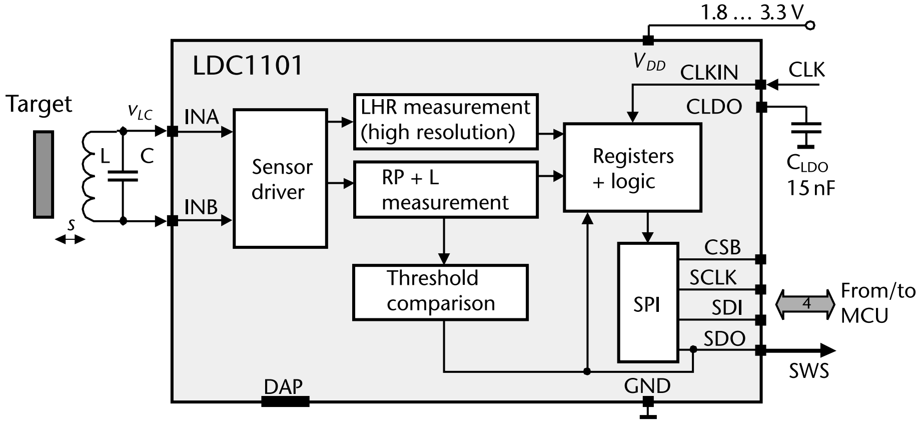

1.8 … 3.3 V
LDC1101
$V_{DD}$
CLKIN
CLK
CLDO
Target
$v_{LC}$
INA
L
C
s
INB
Sensor driver
LHR measurement (high resolution)
RP + L measurement
Registers + logic
$C_{LDO}$ 15 nF
Threshold comparison
SPI
CSB
SCLK
SDI
SDO
4
From/to MCU
SWS
DAP
GND

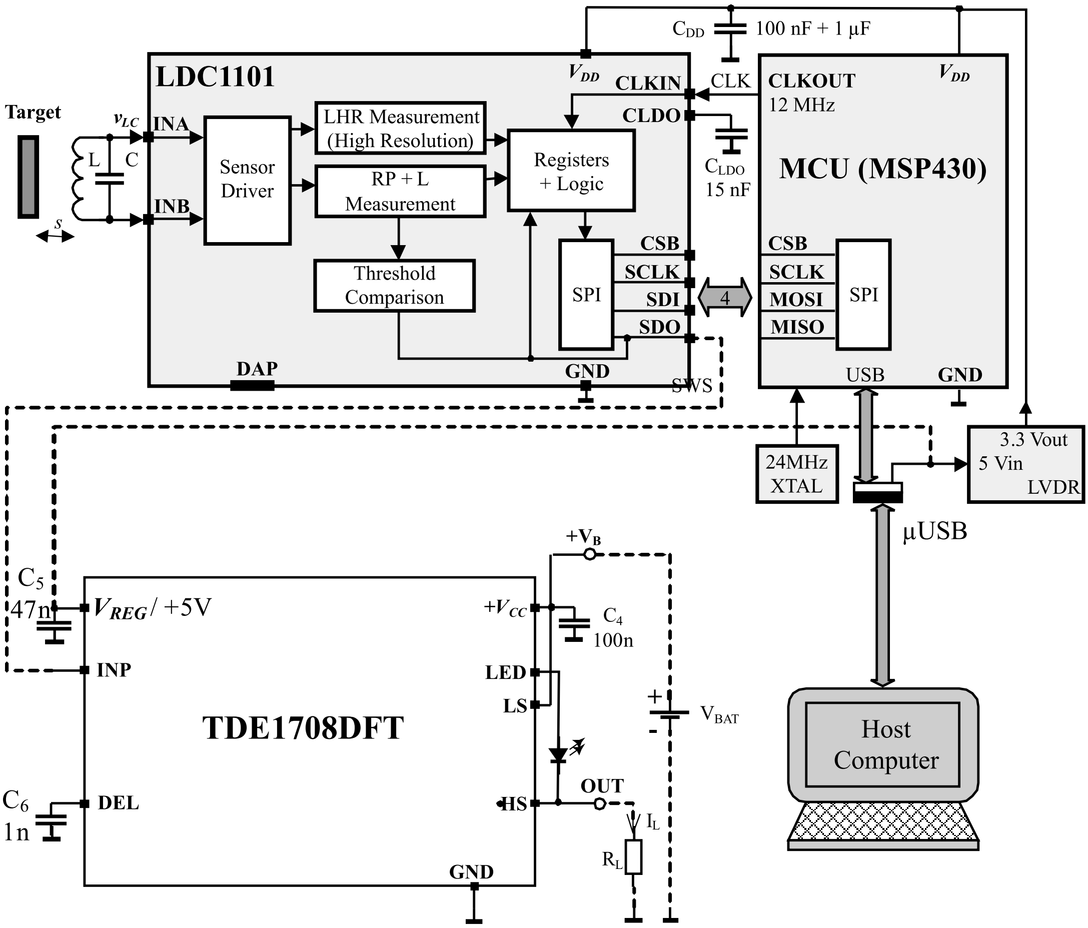
**Figure 7.23** Functional block diagram of the device LDC1101 with manufacturer notation and terminology [26].

#### 7.4.3.2 LDC1101: Typical Application and Evaluation Module EVM

A typical application is shown in Figure 7.24. The LDC1101 is supported by a microcontroller central unit (MCU) as master. The manufacturer recommends a device from the family MSP430x6xx, which unconditionally has the SPI interface. It would be even better if the MCU has a USB option. A 24-MHz quartz crystal XTAL ensures a 12-MHz clock signal for LDC1101. The low-voltage drop regulator (LVDR) supplies both devices with the standard value $V_{DD} = 3.3V$.

If a power amplification of the switching output signal SWS is needed, an extension of the circuitry using, for example, the output driver TDE1708DFT (see Figure 7.21) is opportune. The connection lines to this additional stage and to the standard application are represented in Figure 7.24 by dashed lines.

The communication between LDC1101 and MCU to configure the LDC1101 and for input/output data streams is carried out between these units via the SPI. In addition, a host computer can be used to make the first startup with the initial



LDC1101
Target
L
C
s
v_LC
INA
INB
Sensor Driver
LHR Measurement (High Resolution)
RP + L Measurement
Threshold Comparison
Registers + Logic
SPI
V_DD
CLKIN
CLDO
CSB
SCLK
SDI
SDO
DAP
GND
SWS
C_DD
100 nF + 1 µF
CLK
C_LDO
15 nF
4
MCU (MSP430)
CLKOUT
12 MHz
V_DD
CSB
SCLK
MOSI
MISO
SPI
USB
GND
24MHz XTAL
3.3 Vout
5 Vin
LVDR
µUSB
Host Computer
C5
47n
C6
1n
TDE1708DFT
V_REG / +5V
INP
DEL
+V_CC
LED
LS
HS
GND
+V_B
C4
100n
V_BAT
OUT
I_L
R_L

**Figure 7.24** Typical application of the system with LDC1101 with two options: TI proposal with host computer and author's proposal with power-switching output stage TDE1708DFT, respectively.

configuration easier and for future reconfigurations. Texas Instruments offers a graphical user interface (GUI) free of charge for this application (see the next section).

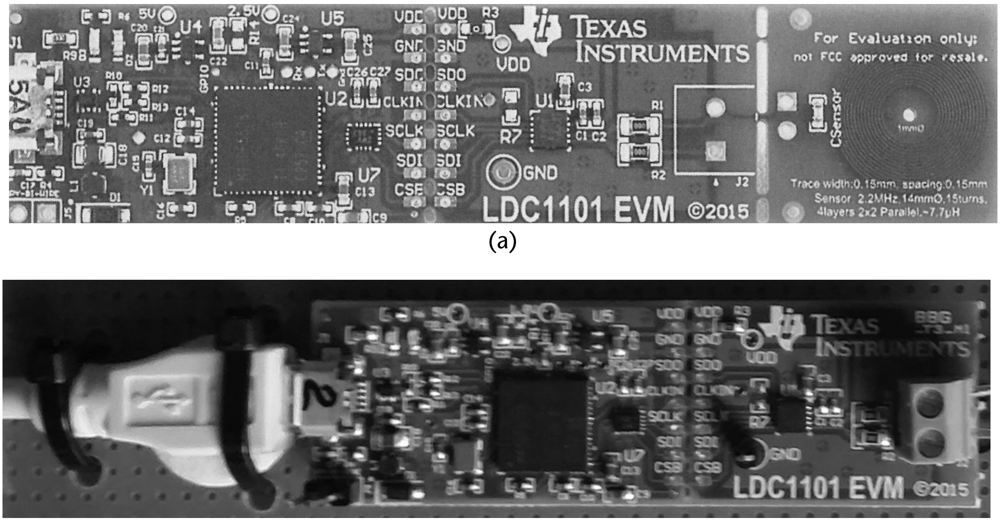
Moreover, the manufacturer gives away samples already assembled on PCB carrier, called LDC1101EVM (evaluation module) [28], which helps users evaluate the operation and programming of LDC1101.

Reference [28] presents the LDC1101EVM schematic, which is similar to the typical application shown in Figure 7.24. The evaluation module is supplied via $\mu$USB connector with 5V and contains a low-cost MCU (MSP430F5528) supplied with 2.5V by means of a voltage regulator. The LDC1101 is supplied with 1.8V by means of a second voltage regulator. The level difference between the connections of MCU and LDC is balanced by a level shifter. The MCU contains preloaded firmware to communicate with LabVIEW-based PC controller software.

From a constructive point of view, the sample's carrier PCB (dimensions 90 mm × 20 mm) is divided in three areas (Figure 7.25(a)). The right section contains a parallel resonant circuit consisting of a spiral flat four-layer PCB coil (see Figure 5.11(a)) with 17 turns (inductance $L = 7.2\ \mu$H) and a 390-pF capacitor. The middle area contains the LDC1101 circuitry and the left area has the MCU, supply stages, and the $\mu$USB connection to the host computer.

Due to the breakaway concept of the PCB, the coil section can be broken along a separation line and removed. A separate other coil and/or resonance capacitor can be reconnected to the LDC1101 (see the screw connector in Figure 7.25(b)). At the limit, the middle section with LDC1101 can also be separated from the left section and can be reconnected to a different MCU platform.

The documentation in [28] describes in great detail how to properly connect, set up, and use the evaluation module LDC1101EVM.


TEXAS INSTRUMENTS
For Evaluation only: not FCC approved for resale.
LDC1101 EVM ©2015
Trace width:0.15mm, spacing:0.15mm Sensor 2.2MHz,14mmØ,15turns, 4layers 2x2 Parallel,~7.7µH
VDD
GND
SDO
CLKIN
SCLK
SDI
CSB
CSensor

(a)

(b)

**Figure 7.25** (a) Evaluation sample module with LDC1101 (company photo) [28], and (b) evaluation sample module with removed PCB coil region and mounted on a testing breadboard (in action).

#### 7.4.3.3 LDC1101: System Programming and GUI

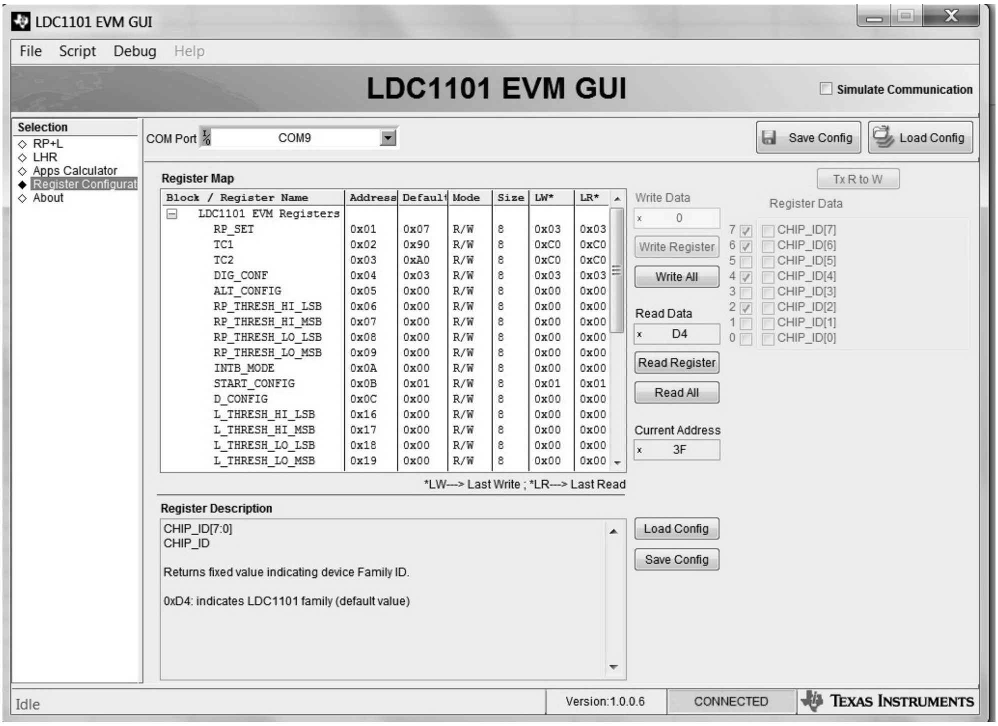
The LDC1101EVM uses PC-based application software for device configuration and to retrieve conversion results. This software is provided by Texas Instruments and can be used to evaluate the system in a simple and flexible GUI configuration.

After successful software installation on the host PC, establishing a USB physical connection between the PC and module, and creating a logical connection by allocation of a virtual COM port, every GUI page should indicate an error-free operation (green CONNECTED field in the right lower corner in Figure 7.26).

The setup is a complex task and requires a previous computation of the values in the configuration registers followed by writing into these registers. Reference [28] gives the computation formulas and how the use them. Then, the registers will be configured by means of read/write facilities of the operating page shown in Figure 7.26.

The application can be started as the GUI page RP+L is selected and the active mode is enabled using the virtual switch Active/Sleep Mode (see the toggle button in the left upper corner in Figure 7.27).

The Running indicator goes green and signals a proper operation of the subsystem *RP + L Measurement*. The $R_P$ and $L$ average values are continuously refreshed in the data outputs registers and forwarded to the master MCU. In parallel, the GUI displays the statistical data of these values: average, minimum, maximum, and standard deviation (see Figure 7.27) and their evolution in real time.



COM Port COM9

Register Map

| Block / Register Name | Address | Defaul | Mode | Size | LW* | LR* |
|---|---|---|---|---|---|---|
| ⊟ LDC1101 EVM Registers | | | | | | |
| RP_SET | 0x01 | 0x07 | R/W | 8 | 0x03 | 0x03 |
| TC1 | 0x02 | 0x90 | R/W | 8 | 0xC0 | 0xC0 |
| TC2 | 0x03 | 0xA0 | R/W | 8 | 0xC0 | 0xC0 |
| DIG_CONF | 0x04 | 0x03 | R/W | 8 | 0x03 | 0x03 |
| ALT_CONFIG | 0x05 | 0x00 | R/W | 8 | 0x00 | 0x00 |
| RP_THRESH_HI_LSB | 0x06 | 0x00 | R/W | 8 | 0x00 | 0x00 |
| RP_THRESH_HI_MSB | 0x07 | 0x00 | R/W | 8 | 0x00 | 0x00 |
| RP_THRESH_LO_LSB | 0x08 | 0x00 | R/W | 8 | 0x00 | 0x00 |
| RP_THRESH_LO_MSB | 0x09 | 0x00 | R/W | 8 | 0x00 | 0x00 |
| INTB_MODE | 0x0A | 0x00 | R/W | 8 | 0x00 | 0x00 |
| START_CONFIG | 0x0B | 0x01 | R/W | 8 | 0x01 | 0x01 |
| D_CONFIG | 0x0C | 0x00 | R/W | 8 | 0x00 | 0x00 |
| L_THRESH_HI_LSB | 0x16 | 0x00 | R/W | 8 | 0x00 | 0x00 |
| L_THRESH_HI_MSB | 0x17 | 0x00 | R/W | 8 | 0x00 | 0x00 |
| L_THRESH_LO_LSB | 0x18 | 0x00 | R/W | 8 | 0x00 | 0x00 |
| L_THRESH_LO_MSB | 0x19 | 0x00 | R/W | 8 | 0x00 | 0x00 |

*LW---> Last Write ; *LR---> Last Read

Write Data x 0

Read Data x D4

Current Address x 3F

Register Data
7 ☑ ☐ CHIP_ID[7]
6 ☑ ☐ CHIP_ID[6]
5 ☐ ☐ CHIP_ID[5]
4 ☑ ☐ CHIP_ID[4]
3 ☐ ☐ CHIP_ID[3]
2 ☑ ☐ CHIP_ID[2]
1 ☐ ☐ CHIP_ID[1]
0 ☐ ☐ CHIP_ID[0]

Register Description

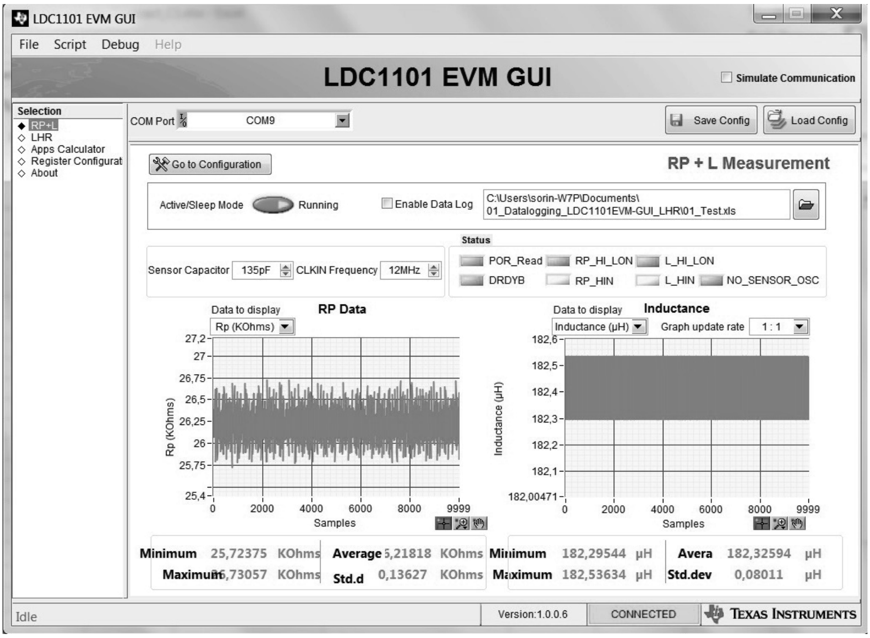
CHIP_ID[7:0]
CHIP_ID

Returns fixed value indicating device Family ID.

0xD4: indicates LDC1101 family (default value)

Idle | Version:1.0.0.6 | CONNECTED

**Figure 7.26** GUI operating page for register configuration.


LDC1101 EVM GUI
File
Script
Debug
Help
LDC1101 EVM GUI
Simulate Communication
Selection
RP+L
LHR
Apps Calculator
Register Configurat
About
Save Config
Load Config
Go to Configuration
RP + L Measurement
Status
POR_Read
RP_HI_LON
L_HI_LON
DRDYB
RP_HIN
L_HIN
NO_SENSOR_OSC
RP Data
Inductance
Rp (KOhms)
27,2
27
26,75
26,5
26,25
26
25,75
25,4
0
2000
4000
6000
8000
9999
Samples
Inductance (µH)
182,6
182,5
182,4
182,3
182,2
182,1
182,00471
Texas Instruments

☐ Simulate Communication

COM Port: COM9

Active/Sleep Mode: Running

☐ Enable Data Log: C:\Users\sorin-W7P\Documents\01_Datalogging_LDC1101EVM-GUI_LHR\01_Test.xls

Sensor Capacitor: 135pF
CLKIN Frequency: 12MHz

Data to display: Rp (KOhms)

Data to display: Inductance (µH)
Graph update rate: 1 : 1

Minimum 25,72375 KOhms
Maximum 26,73057 KOhms
Average 26,21818 KOhms
Std.d 0,13627 KOhms

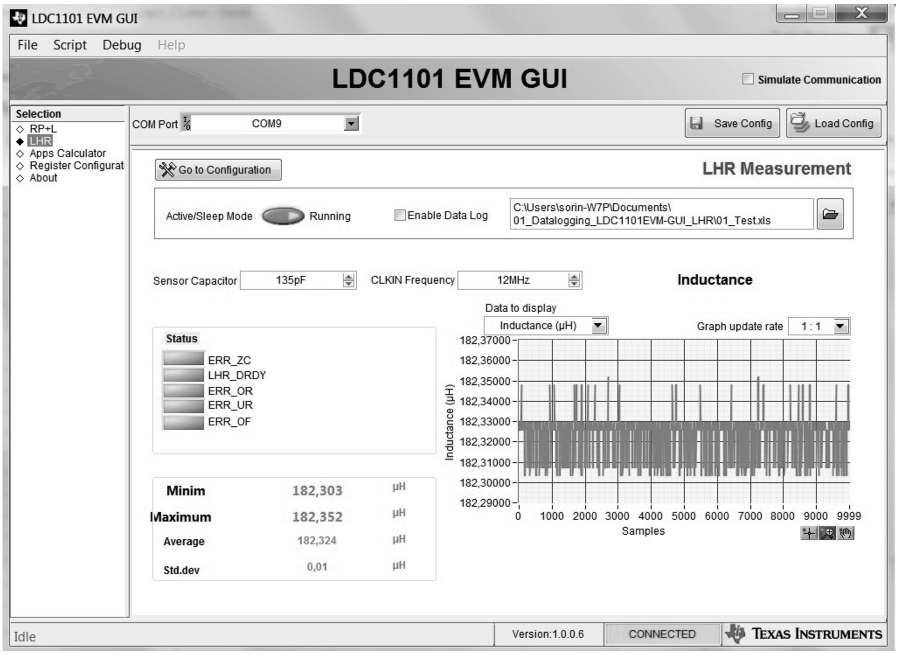
Minimum 182,29544 µH
Maximum 182,53634 µH
Average 182,32594 µH
Std.dev 0,08011 µH

Idle
Version:1.0.0.6
CONNECTED

**Figure 7.27** GUI operating page for RP + L measurement.

As mentioned, the LHR measurement runs in parallel and the results are available in a similar manner (see Figure 7.28).

For every change of the register configuration it is necessary to enable the sleep mode, to operate the changes, and to reenter in the active mode.


LDC1101 EVM GUI
File
Script
Debug
Help
LDC1101 EVM GUI
Simulate Communication
Selection
RP+L
LHR
Apps Calculator
Register Configurat
About
Save Config
Load Config
Go to Configuration
LHR Measurement
Inductance
Status
ERR_ZC
LHR_DRDY
ERR_OR
ERR_UR
ERR_OF
Inductance (µH)
182,37000
182,36000
182,35000
182,34000
182,33000
182,32000
182,31000
182,30000
182,29000
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
9999
Samples
Texas Instruments

☐ Simulate Communication

COM Port: COM9

Active/Sleep Mode: Running

☐ Enable Data Log: C:\Users\sorin-W7P\Documents\01_Datalogging_LDC1101EVM-GUI_LHR\01_Test.xls

Sensor Capacitor: 135pF
CLKIN Frequency: 12MHz

Data to display: Inductance (µH)
Graph update rate: 1 : 1

Minim 182,303 µH
Maximum 182,352 µH
Average 182,324 µH
Std.dev 0,01 µH

Idle
Version:1.0.0.6
CONNECTED

**Figure 7.28** GUI operating page for LHR measurement.

*Example 7.1*

To illustrate the GUI features, a very common ISE of a popular inductive proximity sensor was connected and measured by means of an LDC1101EVM. Its design parameters were $L = 188\ \mu\text{H}$ and $C = 135$ pF.

Critical to the system configuration are the calculated contents: 0 × 03, 0 × C0, 0 × C0, and 0 × 03 of the registers 0 × 01 to 0 × 04, respectively. As described, they have been stored in registers and are visible on the GUI page in Figure 7.26.

The $RP + L$ subsystem measured and displayed average values of $R_P = 25.21$ kΩ and $L = 182.32\ \mu\text{H}$ (see Figure 7.27).

The LHR subsystem measured and provided an average value of the inductance $L = 183.30\ \mu\text{H}$ (see Figure 7.28).

Concluding the presentation of LDC1101, it is appropriate to call this device an application-specific intelligent sensor (ASIS), even though this term is used far less frequently than ASIC.

## 7.5 Software-Defined Sensors: Fantasy or the Inductive Sensor of Tomorrow?

To speak today about software-defined sensors (SDS) is perhaps a bit too visionary. It is quite similar to exploring the future.

SDS is at the moment only a concept. It would then be only a question of time before SDS becomes reality. The hardware and software technical requirements are at this stage available; only the questions regarding the costs and dimensions are open.

The target setting of SDS is to replace classical analog stages in the signal processing chain (see Figure 7.1), if possible, by a single digital device (all-purpose microcontroller) as well as to develop and preload firmware with control subprograms and algorithm subroutines. This hardware/software platform should provide an entire sensor functionality: digitalizing the primary signal from the sensing element (sampling), full digital signal processing, generation of the digital output signals, and optionally, reconversion of some final digital signals into analog output signals.

The sensor manufacturers struggle to obtain legal rights by protected trademark registration [29]. The application characteristics are hereby broadly defined: "SDS is a system for tuning, analysing performance and/or diagnosing electric or electronic sensors and/or creating new electric or electronic sensors; sensor architecture comprised of hardware sensor and the software layer to modify functions and performance and create new sensors at optimized power profiles."

The basic concept of SDS is that a sensor could be totally configured or defined by the software so that a common hardware support can be used across a number of sensor types. In parallel, the software is used to change the configuration for functions required at a given time. The sensor can be reconfigured as an upgrade or, if required, to meet changing demands.

Although it may sound like a trivial exercise, creating a definition for SDS is not as simple as it seems. Theoretically speaking, an SDS is a sensor in which some or all of the physical layer functions are software-defined. In other words, the software is used to determine how and what the sensor does.

A more practical definition is that an SDS is a system that has a generic hardware platform on which software runs to provide specific functions, as shown in Figure 7.1 (oscillator, demodulator, etc.). By reconfiguring the software, the sensor functionality is changed. To achieve this goal, SDS technology should use software modules (signal generation, signal analyzing, etc.) that run on a generic hardware platform consisting of general-purpose processors or digital signal processing (DSP) processors.

It is not always feasible or practicable to develop a sensor that incorporates all the features of a full SDS. Some sensors may only support a number of features associated with SDSs whereas others may be fully software-defined. For this reason, it makes sense to define more SDS levels.

Ideally, the signal from the primary sensing element would be directly converted to digits and all processing steps up to the sensor output be undertaken under software control. In this way, there are no limitations introduced by the hardware. To achieve this, the analog-to-digital and digital-to-analog conversion at the sensor input and output, respectively, would need to have a relatively high working frequency, accuracy, and resolution.

This is the highest SDS level 3 where the boundaries between configurable and nonconfigurable areas are very close to the sensing element and to the application, respectively. It could be considered to have full programmability.

On the SDS level 2, a significant proportion of the sensor architecture is software-configurable. Parameters and features are software-controlled. The front-end and the back-end still partially remain hardware-based and nonreconfigurable.

In terms of what is configurable, this approach leads to the assumption of a subordinate level 1, where limited functions are controllable by software. A good example is the system with LDC1101 in Figure 7.24.

Finally, an inferior level 0 could be also defined for the nonconfigurable hardware sensor (i.e., one that cannot be changed by software, such as the system with TCA505 shown in Figure 7.17).

The block diagram in Figure 7.29 illustrates the SDS level 2. If the stages inside the range bordered by dashed line are bundled in a single device, it is possible to speak about the ideal SDS level 3.

The output signal(s) of the ISE will immediately be digitalized. The Nyquist-Shannon sampling theorem requests a sampling frequency being two times or higher than the frequency of the ISE output. This requirement is comfortably fulfilled by the up-to-date 12-bit A/C convertors.

The core of the SDS, the digital signal processor, determines the power of the sensor. Every IC supplier of note has at least one such device in their portfolio.

A suitable device for SDS implementation could be offered by Microchip Technology Inc. from the wide spectrum within the series of dsPIC. These DSP-controllers with Harvard-architecture are characterized by the following main performances [30]:

- Strong DSP engine (16 × 16 bit multiplier, two 40-bit accumulators);
- Two 12-bit A/D convertors, clock generator, timers, etc.;
- High-capacity RAM and FRAM memories (alternatively EEPROM);
- Rich periphery: SPI, I$^2$C, UART, and CAN interfaces (for details see Chapter 14);

- Clock frequency ≤30 MHz, supply voltage 2.5 to 5V;
- Code efficient (C and assembly) architecture.

An alternative solution is a microcontroller TM320F Piccolo™ from the Texas Instruments [21] series having the same Harvard architecture.

Comparable applicability is offered by the processor core family MIPS32®microAdaptiv™ (supplier www.imgtec.com), which in particular has a microprocessor without interlocked pipeline stages (MIPS architecture).

The preloaded firmware in the SDS system has essentially two major duties: (1) to provide the sensor base operation and (2) to support the signal processing with a high algorithmic quality in order to achieve the needed sensor intelligence. The scope of intelligence is very comprehensive. The result of a survey of sensor product managers a few years ago led to the following surprising ranking list of sensor intelligence attributes: diagnosis possibility (top position), plug-and-play, tech-in, adaptability and configuration attributes, multisensor options, high sensitivity, accuracy and resolution, high-grade processing using effective up-to-date algorithms (low-ranking position). This is the market opinion.

From the view of intelligent signal recognition and processing, the systems presented in the previous sections of this chapter unilaterally perform a signal analysis. They play the role of analyzer in a classical pattern recognition system [31]. Sensor intelligence can be significantly increased cascading a follower subsystem after the analyzer in an advanced system, namely the vector classifier [31]. Suitable classifier types are neuronal networks [32], fuzzy classifiers, and support vector machines.

The signal processing chain in Figure 7.29 ends with analog as well as binary and digital sensor outputs (see Section 1.3.7). The intermediate binary SDS outputs and the D/A outputs need to be driven to become standard-compliant. According to the above-mentioned survey, the ranking list of the future SDSs outputs should be the following:

- Standard outputs: analog and binary;
- Wireline standard IO-Link (for details see Chapter 14);
- Wireless standard IO-Link;
- Internet (LAN/WLAN).

The additional signals path, environment transducers and capturing, is necessary to inform the system about environmental factors with negative influence (first being the temperature) in order to compensate them.

The desire for use of energy harvesting sources such as mechanical energy resulting from effects such as vibration, stress and strain, and thermal energy is not easy to fulfill at present, because the DSP devices need significant supply energy. A classical supply with energy from a power network is currently more realistic.

Looking forward, there are several reasons for the need for SDSs:

- *Cost savings:* There is a real need to be able to reuse available platforms for very different projects.
- *Obsolescence mitigation:* A similar requirement comes as hardware technology develops and it is necessary to transfer existing functionalities onto newer platforms.

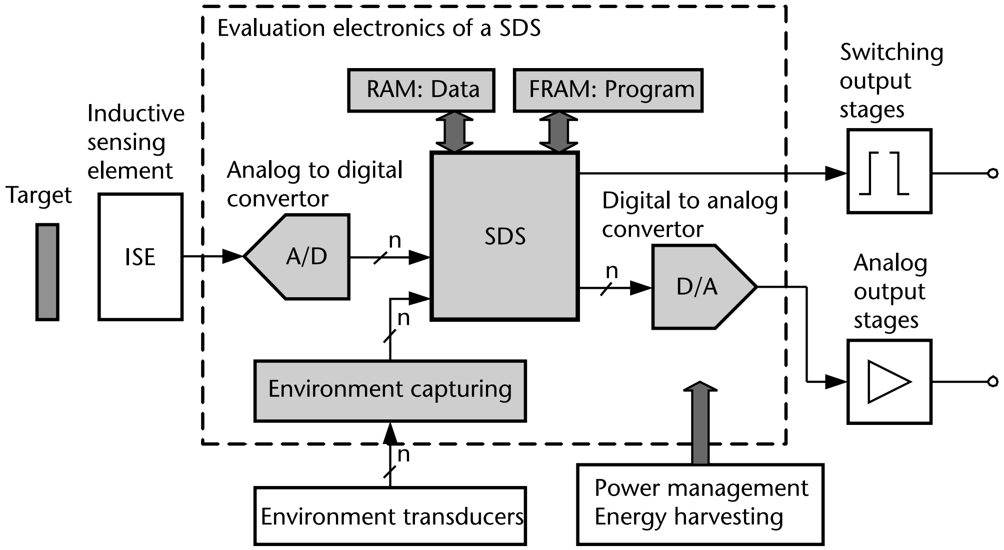

Evaluation electronics of a SDS
RAM: Data
FRAM: Program
Switching output stages
Inductive sensing element
Target
ISE
Analog to digital convertor
A/D
n
SDS
Digital to analog convertor
n
D/A
Analog output stages
n
Environment capturing
n
Environment transducers
Power management
Energy harvesting

**Figure 7.29** Block diagram of an SDS.

- *Interoperability:* To provide complete interoperability a customer may request the use of a particular functionality being used across the equipment from several manufacturers.
- Last but not least, the major *technological benefit* consists of the possibility to greatly increase sensor intelligence. Complex measuring jobs and operations, which lead to significant improvement of sensor features, better stability even by a higher sensibility, better compensation of unwanted environmental influences, and so forth, are only possible using high-level algorithms and up-to-date evaluation methods. They need very high computation power and large hardware resources of memory, interfaces, speed, and so forth, that are specific to DSP processors.

In conclusion, it should be stated that the information relative to ASICs for inductive sensors in this chapter, as well as in the following chapters, is presented on a large scale based on the author's knowledge of this field. This experience was gained over 25 years, in cooperation with ASIC projects done by six major organizations around the world: design houses and well-known IC companies (with fab or fabless), including Texas Instruments Incorporated (two ASIC projects).

The author was involved with, managed, and actively participated as a beneficiary and/or designer at the concept elaboration, creating specifications, samples tests and evaluation, improvements and redesign, final testing, and acceptance of the serial devices.

## References

[1] Tietze, U., C. Schenk, and E. Gamm, *Halbleiter-Schaltungstechnik*, 15th Edition, Berlin Heidelberg, Germany: Spinger/Vieweg-Verlag, 2016.

[2] Langhein, W., *Annaeherungsschalter*, Elektrotechnische Zeitschrift, Band 4, Heft 9, 1952.

[3] https://en.wikipedia.org/wiki/Triode.

[4] https://spectrum.ieee.org/tech-history/silicon-revolution/the-lost-history-of-the-transistor.

[5] Amos, S., and M. James, *Principles of Transistor Circuits*, E-book, 9th Edition, Oxford, UK: Newnes, 2003.

[6] Burghartz, J. N. (ed.), *Guide to State-of-the-Art Electron Devices*, Chichester, UK: John Wiley & Sons, 2013.

[7] Gray, P. R., P. J. Hurst, H. S. Lewis, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits, Fifth Edition*, New York: John Wiley & Sons, 2009.

[8] Daly, J. C., and D. P. Galipeau, *BiCMOS Design: Practices and Pitfalls*, Boca Raton, FL: CRC Press, 2000.

[9] Lau, J. H. (ed.), *Flip Chip Technologies*, Boston: McGraw-Hill Professional, 1995.

[10] Lau, J. H., R. S. W. Lee, *Chip Scale Package: Design, Materials, Process, Reliability, and Applications*, New York: McGraw-Hill Professional, 1999.

[11] www.xfab.com, XC06_CMOS Process, 0.6 Micron Modular Mixed Signal Technology, Erfurt, Germany and Lubbock, TX, 2018.

[12] www.vishay.com/doc70195, *CR160 Series Current Regulator Diodes–Vishay*, PDF document, 2004.

[13] https://www.murata.com, *Application Manual PTC Thermistor POSISTOR*, 2014.

[14] http://www.ti.com/lit/ds/symlink/lm340.pdf, *Data Sheet: LM340 and LM7805 Family, Wide VIN and 1.5 A Fixed Regulators*, 2016.

[15] https://www.infineon.com/dgdl/tca505_green_050217.pdf, *Datasheet TCA505BG*, 2005.

[16] Siemens AG, *Product Information Booklet: ICs for Inductive Proximity Switches*, 2005.

[17] Hoedlmayr, F., *TCA 505: Naeherunsschalter-IC mit Kurzschlussschutz*, Siemens Components 28, Heft 5, 1990.

[18] Herfurth, M., IC fuer induktive Naeherunsschalter-Eigenschaften, Der Elektroniker, Nr.5/1991.

[19] www.elektroniknet.de, *Maschinen-Sensor*: *Naeherungsschaltern mit zwei ICs aufgebaut*, 2009.

[20] www.st.com, *STMicroelectronics*–Product Catalog, 2018.

[21] www.ti.com/products, *TI Analog, DSP and Semiconductor Products*, Texas Instruments, 2017.

[22] Jagiella, M., S. Fericean, and A. Dorneich, "Progress and Recent Realizations of Miniaturized Inductive Proximity Sensors for Automation," *Sensors Journal*, Vol. 6, No. 6, 2006.

[23] Fericean, S., and R. Droxler, "New Noncontacting Inductive Analog Proximity and Inductive Linear Displacement Sensors for Industrial Automation," *IEEE Sensors Journal*, Vol. 7, No. 11, 2007.

[24] https://www.analog-micro.com, *AM401 Data Sheet*, Analog Microelectronics GmbH. 2006.

[25] Fericean, S., H. Kammerer, and H.- W. Plank, *Proximity Switch Operating in a Non-Contacting Manner*, US Patent No. 5,408,132, issued April 18, 1995.

[26] www.ti.com/lit/ds/symlink/ldc1101.pdf, *LDC1101 1.8-V High-Resolution, High-Speed Inductance-to-Digital Convertor*, Texas Instruments, October 2016.

[27] https://www.corelis.com/education/SPI_Tutorial.htm, *SPI Interface–Tutorial*, Corelis, 2018.

[28] www.ti.com/lit/ug/snou137/snou137.pdf, *LDC1101EVM User's Guide*, Texas Instruments, May 2015.

[29] www.euipo.europa.eu, European Intellectual Property Office, *014192744—Balluff Software Defined Sensor*, 08/102015.

[30] www.microchip.com, *16-Bit Microcontrollers and Digital Signal Controllers dsPIC33*, datasheet, 2013.

[31] Hoffmann R., and M. Wolff, *Intelligente Signalverarbeitung 2–Signalerkennung*, 2nd Edition, Berlin, Germany: Spinger/Vieweg-Verlag, 2015.

[32] Polik, Z., and Kantor, Z., *Inductive Displacement Measuring Sensor and Method for Operating the Latter*, U.S. Patent Application 14/778,219, filed March 17, 2016.

## CHAPTER 8

# Excitation and Evaluation of the Inductive Sensing Element: Oscillators

An oscillator is the core part of an inductive sensor; it exists in every sensor and generates the working frequency of this system. For this purpose, the oscillator contains a resonant element, which is a resonant circuit (coil and capacitor) in the majority of cases. A small number of sensors contain other resonant elements like ceramic or quartz components.

The coil can be influenced by the target or not. If it is, the oscillator plays a double role: it delivers the working frequency and supplies the inductive sensing element ISE with AC current but at the same time performs the primary target detection. If the inductor is not influenced by the target, the oscillator delivers a driving signal, whose frequency is generally constant and has to supply the inductive sensing element consisting of an additional detection coil.

## 8.1 Theory of Resonant LC Circuits: Series versus Parallel

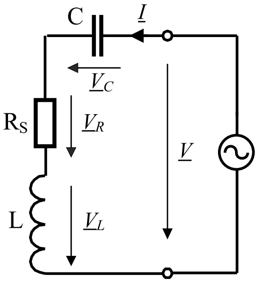
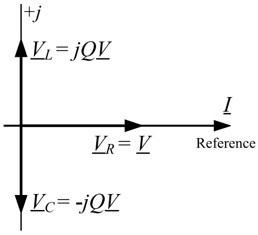
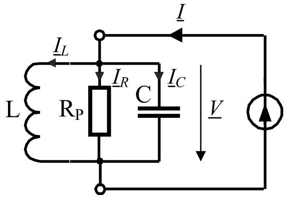
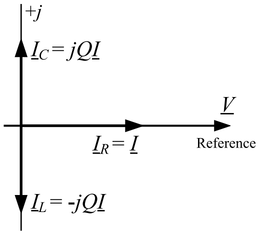
Resonant circuits were introduced in Chapter 4. After theoretical considerations in Sections 4.1.1 and 4.1.2 relative to the inductor as a key component of the inductive sensing elements (e.g., definitions of its inductance, impedance, and quality factor), Section 4.1.3 defines the two classical topologies of the resonant circuits: series and parallel (see Figure 4.6) and gives formulas for their impedance and resonance frequency. Sections pay attention to the details of the parallel version, which is more popular for inductive sensors (see Table 4.1). At this point, Chapter 4 ends the description of the resonant circuits.

This section revisits resonant circuits and provides benchmarking from other points of view, which are significant for the theory of electronic oscillators. For a better overview, the main features are summarized in Table 8.1.

A preliminary assumption is that circuits are not ideal and have losses expressed by a series loss resistance $R_S$ or a parallel resistance $R_P$ (see (4.38)). Applying a driving signal to a resonant circuit can overcome these losses of the circuit and cause the resonant circuit to oscillate indefinitely [1].

For informative purposes only, the computed features and graphical representations in Sections 8.1.1 and 8.1.2 are done for specific resonant circuits having $L = 1$ mH, $C = 1$ nF, $R_S = 30\Omega$, or $R_P = 33.33$ k$\Omega$, respectively.

**Table 8.1** Benchmarking between the Series and Parallel Resonant LC Circuit

| *Attribute* | *Series resonant LC-circuit (SRC)* | *Parallel resonant LC-circuit (PRC)* |
|---|---|---|
| *Schematics*<br>*Phasor Diagrams at Resonance* |  C<br>$\underline{I}$<br>$\underline{V}_C$<br>$R_S$<br>$\underline{V}_R$<br>$\underline{V}$<br>L<br>$\underline{V}_L$<br> +j<br>$\underline{V}_L = jQ\underline{V}$<br>$\underline{I}$<br>$\underline{V}_R = \underline{V}$<br>Reference<br>$\underline{V}_C = -jQ\underline{V}$ |  $\underline{I}$<br>$\underline{I}_L$<br>$\underline{I}_R$<br>$\underline{I}_C$<br>C<br>L<br>$R_P$<br>$\underline{V}$<br> +j<br>$\underline{I}_C = jQ\underline{I}$<br>$\underline{V}$<br>$\underline{I}_R = \underline{I}$<br>Reference<br>$\underline{I}_L = -jQ\underline{I}$ |
| *Resonance Frequency* | $f_0 \approx \frac{1}{2\pi} \cdot \sqrt{\frac{1}{LC}}$ | $f_0 \approx \frac{1}{2\pi} \cdot \sqrt{\frac{1}{LC}}$ |
| *Q-factor* | $Q = \frac{1}{R_S} \cdot \sqrt{\frac{L}{C}}$ | $Q = R_P \cdot \sqrt{\frac{C}{L}}$ |
| *Attenuation Factor* | $d \overset{\text{def}}{=} \frac{1}{Q} = R_S \cdot \sqrt{\frac{C}{L}}$ | $d \overset{\text{def}}{=} \frac{1}{Q} = \frac{1}{R_P} \cdot \sqrt{\frac{L}{C}}$ |
| *−3 dB-Bandwidth [see (4.39)]* | $B_Q \overset{\text{def}}{=} \frac{f_0}{Q} = d \cdot f_0 = \frac{R_S}{2\pi L}$ | $B_Q \overset{\text{def}}{=} \frac{f_0}{Q} = d \cdot f_0 = \frac{1}{R_P} \cdot \frac{1}{2\pi C}$ |
| *Kind of resonance* | *Voltage Resonance* | *Current Resonance* |

### 8.1.1 Characteristics of the Series Resonant LC Circuit

The behavior of the series resonant LC circuit (SRC) can be summarized as follows:

1. Below the resonance frequency $f_0$ (≈159 kHz in our numerical example), the magnitude of the negative capacitive reactance $X_C = -1/\omega C$ is higher than the magnitude of the positive inductive reactance $X_L = \omega L$, the sum $X_L + X_C$ (series circuit) has a negative sign, and the SRC behavior is capacitive-resistive (Figure 8.1).
2. Above the resonance frequency $f_0$, the sum changes its sign and the positive sign denotes an inductive-resistive behavior of the SRC. Due to the nonlinear component $X_C$, the slope of the curve of impedance magnitude $|\underline{Z}|$ is lower in this frequency range (see (4.28)).
3. At the resonance frequency $f_0$:

- The magnitudes of the reactances become equal: $|X_L| = |X_C|$. They cancel each other out and the SRC impedance becomes a real quantity equal to $R_S$ (Figure 8.1).
- The voltages over the inductor and the capacitor lead/leg the supply current by 90 degrees, respectively. In the phasor diagram of the SRC (Table 8.1), the corresponding complex representations (phasors $\underline{V}_L$ and $\underline{V}_C$) have vertical orientation along the axis of ordinate.
- They have a multiplied magnitude (by Q factor) relative to the supply voltage $\underline{V}$ that denotes the specific voltage resonance that characterizes the series resonant circuit. The voltages at the terminals of the capacitor and of the inductor are by a factor of Q higher than the supply voltage.
- The voltage $v_R$ is in phase with the current $i$ (reference).
- The current through SRC achieves a peak and a zero phase angle (Figure 8.2).

### 8.1.2 Characteristics of the Parallel Resonant LC Circuit

The behavior of the parallel resonant LC circuit (PRC) is optimally analyzed using the admittance $\underline{Y}$ with its components: conductance $G$ and susceptance $B$ (see (4.13)) instead of the impedance $\underline{Z}$. It is expressed by:

$$\underline{Y} \overset{\text{def}}{=} \frac{1}{\underline{Z}} = G + jB = \frac{1}{R_P} + \frac{1}{X_L} + \frac{1}{X_C} = \frac{1}{R_P} + j\left(\omega C - \frac{1}{\omega L}\right) \tag{8.1}$$

The benefit of admittance is named in Section 4.1.1 and consists of its better adequateness for parallel networks: admittance of the parallel network is simply the sum of the partial admittances. The representation in Figure 8.3 for the PRC is

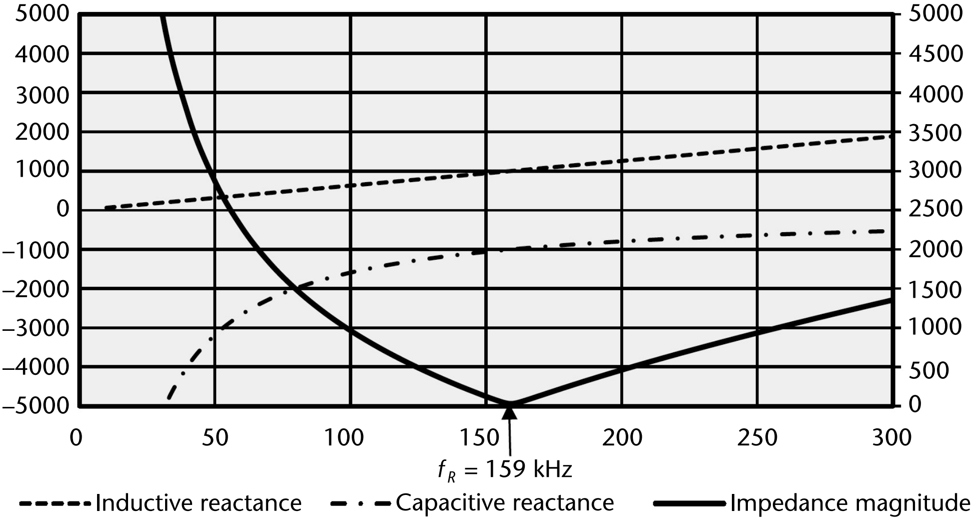


**Figure 8.1** Frequency sweep of the SRC parameters: impedance magnitude $|\underline{Z}|$ (right Y-axis) as well as inductive and capacitive reactances $X_L$ and $X_C$ (left Y-axis).

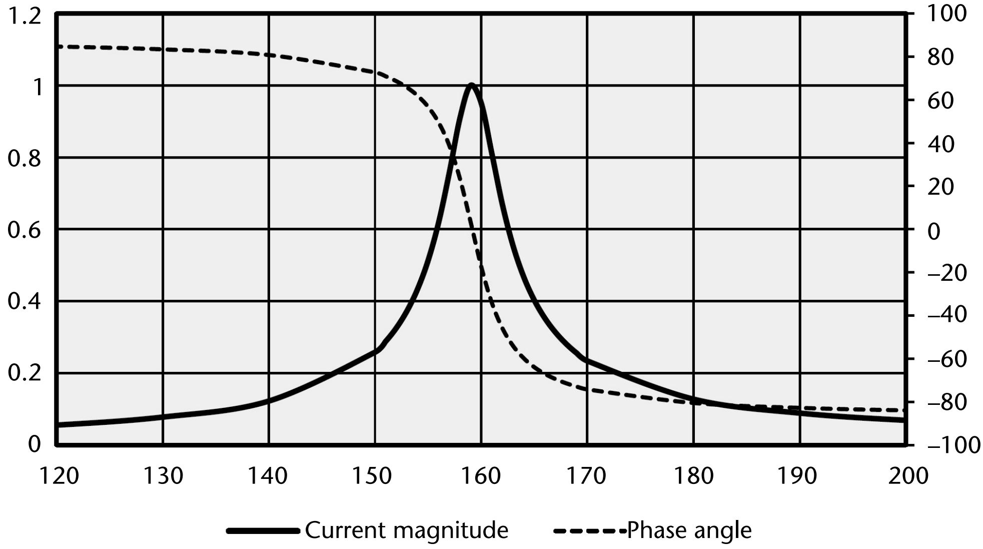


**Figure 8.2** Frequency sweep of the SRC parameters: normalized current magnitude (left *Y*-axis) and phase angle (right *Y*-axis).

similar to the representation in Figure 8.1 for the SRC although it shows the components of the admittance $\underline{Y}$.

The behavior of the PRC can be summarized as follows:

1. Below the resonance frequency $f_0$ ($\approx$ 159 kHz in our numerical example), the magnitude of the negative inductive susceptance $B_L = -1/\omega L$ is higher than the magnitude of the positive capacitive susceptance $B_C = \omega C$, the sum $B_C + B_L$ has a negative sign, and the PRC behavior is inductive-resistive (Figure 8.3).

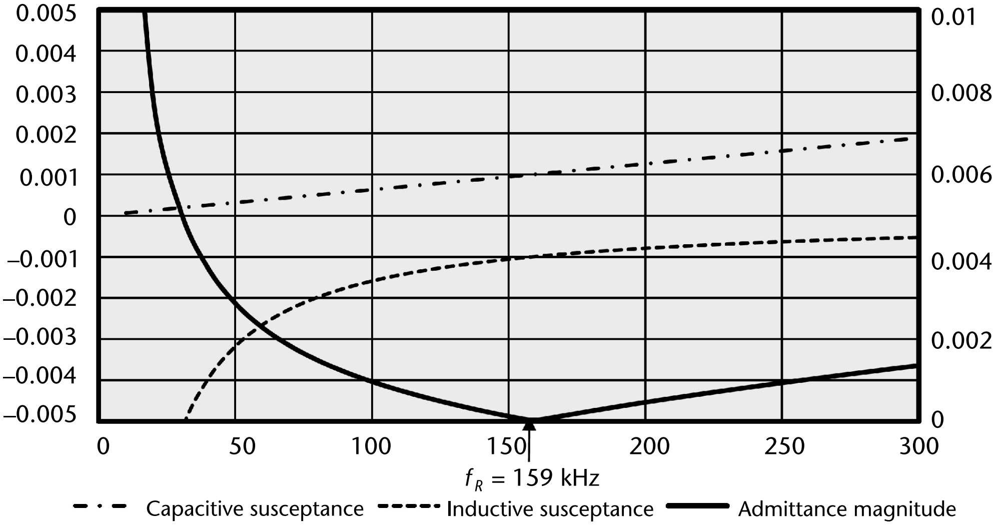


**Figure 8.3** Frequency sweep of the PRC parameters: admittance magnitude $|\underline{Y}|$ (right *Y*-axis) as well as capacitive and inductive susceptances $B_C$ and $B_L$ (left *Y*-axis).

2. Above the resonance frequency $f_0$, the sum changes its sign and the positive sign denotes a capacitive-resistive behavior of the PRC.
3. At the resonance frequency $f_0$
   - The magnitudes of the susceptances become equal: $|B_C| = |B_L|$. They cancel each other out and the PRC admittance becomes a real quantity equal to $1/R_P = 0.03$ mS (Figure 8.3).
   - The currents through the capacitor and the inductor lead/leg the supply voltage by 90 degrees, respectively. In the phasor diagram of the PRC (Table 8.1), the corresponding complex representations (phasors $\underline{I}_L$ and $\underline{I}_C$) have vertical orientation along the axis of ordinate.
   - They have a multiplied magnitude (by Q factor) relative to the supply current $\underline{I}$ that denotes the specific current resonance that characterizes the parallel resonant circuit. The currents through the capacitor and the inductor are by a factor of Q higher than the supply current. This is an important benefit for the inductive proximity sensors: the current through the sensing coil is by factor of Q higher than the excitation current and causes an H-field amplification.
   - The current $i_R$ is in phase with the voltage $v$ (reference).
   - The voltage over the PRC achieves a peak and a zero phase angle (Figure 8.4).

The work with phasor representation is a useful tool to describe the behavior of the resonant LC circuits. Figure 8.5 presents this representation for a PRC, which is supplied with a current having the frequency equal to the resonance frequency $f_0$, but also frequencies lower/higher than this characteristic value.

The vectorial sum of the partial current phasors gives a total current $\underline{I}$, which legs/leads the voltage $\underline{V}$ over the circuit, respectively. The magnitude and the phase

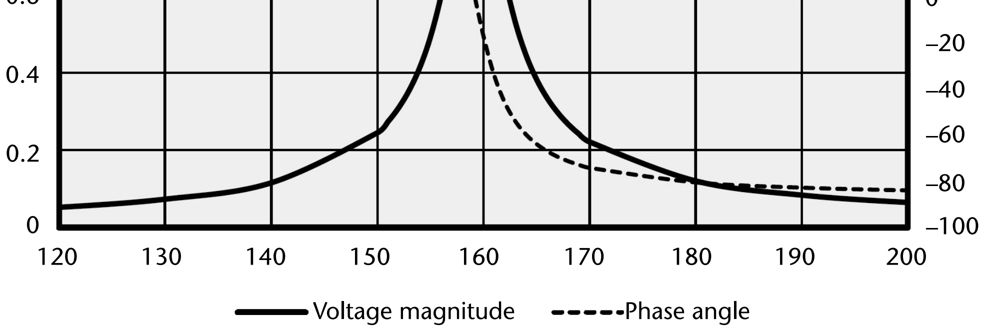


**Figure 8.4** Frequency sweep of the PRC parameters: normalized voltage magnitude (left Y-axis) and phase angle (right Y-axis).

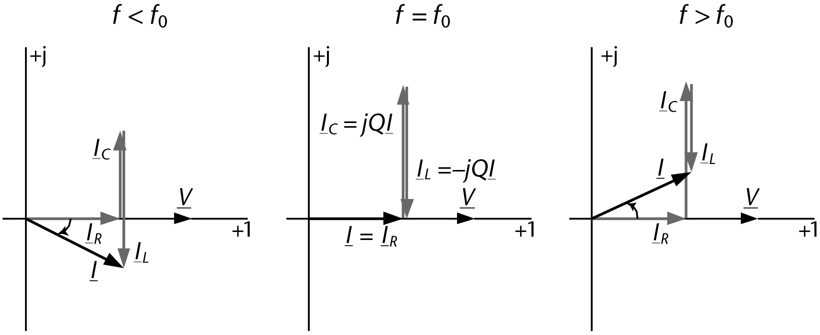


**Figure 8.5** Phasor diagrams of the parallel resonant LC circuit for three different frequencies of the supply signal.

angle of the total current can be graphically determined. In addition, the angle relationship between the total current $\underline{I}$ and the voltage $\underline{V}$ over the circuit denotes the circuit behavior as an inductive or capacitive element, respectively. Obviously, the tool can also be used for SRCs.

Finally, note [2], which makes—in the author's opinion—the most broadly based study of the PRC according to its importance and widespread application in electronic oscillators, and gives additional information over and above what is included here.

## 8.2 General Theory of the Oscillator

The oscillator is presented and described in many publications, although the theory of oscillators is viewed from different perspectives. The author's aspiration was to achieve a systematic summary that joins the most popular study techniques adequate for inductive sensor oscillators.

The comprehensive and in-depth reference work [3] introduces the electronic oscillators by analogy with *mechanical resonators* and describes a large number of resonant systems.

Mechanical resonators work by transferring energy repeatedly from kinetic to potential form and back again. In a pendulum, for example, all the energy is stored as gravitational energy (a form of potential energy) when the bob is instantaneously motionless at the top of its swing. This energy is proportional to both the mass of the bob and its height above the lowest point. As the bob descends and picks up speed, its potential energy is gradually converted to kinetic energy (energy of movement), which is proportional to the bob's mass and to the square of its speed. When the bob is at the bottom of its travel, it has maximum kinetic energy and minimum potential energy. The same process then happens in reverse as the bob climbs toward the top of its swing.

Based on the type of equation, which constructs a mathematical model of the physical system, [3] proposes the following division of oscillatory systems into

- Linear and nonlinear;
- Conservative and nonconservative;
- Self-oscillatory;
- Dissipative, and others

Furthermore, for linear oscillatory systems, in which the motion is described by linear differential equations, [3] names four fundamental classes that basically correspond to the adequate oscillators for inductive sensors:

1. Harmonic oscillators;
2. Oscillators with small mass;
3. Linear systems with negative friction;
4. Linear systems with repulsive force

### 8.2.1 Harmonic Oscillator

The systems called harmonic oscillators have a physical model described by a linear differential equation in the form:

$$\ddot{x} + \omega_0^2 x = 0 \tag{8.2}$$

where $x$ is the system output, $\ddot{x}$ is the second derivative of $x$ with respect to time, and $\omega_0$ is the pulsatance (angular frequency) of the system.

An example of such a system is a body of mass $m$, which horizontally moves along a rod under the action of two lateral springs. Making two simplifying assumptions, namely that (1) the spring force on the body is proportional to its displacement $x$ with respect to its position of equilibrium and (2) the motion is not subject to any friction (caused by air, road, spring), the system equation is:

$$m\ddot{x} + kx = 0 \tag{8.3}$$

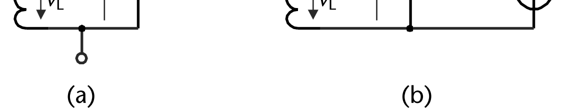
where $k$ is the spring constant.

The replacement $k/m = \omega_0^2$ yields (8.3).

An oscillating circuit consisting of a capacitor C and an inductor L (Figure 8.6(a)), also known as a Thomsonian circuit, is an analogous electrical system.

Initially, in order to arrive at the case of a linear system without friction it is suitable to idealize the circuit assuming that (1) no loss of energy occurs in the system (no connecting leads resistance, no dissipations in the coil conductor or capacitor dielectric and no radiation of the electromagnetic energy), and (2) the capacitance $C$ of the capacitor does *not* depend on its charge nor the inductance $L$ of the inductor on the current intensity through it.

The system equation is determined making two actions in the model shown in Figure 8.6(b):



(a) (b)

**Figure 8.6** (a) Ideal harmonic LC oscillator, and (b) modeling of the establishment of an oscillating state.

1. Defining the initial conditions: The switch $SW_2$ remains open, while the switch $SW_1$ will be switched on. The capacitor C is linearly charged from the ideal DC current source $I$ and after a time interval $t_1$ its voltage $v_C$ is given by:

$$v_C \overset{\text{def}}{=} \frac{q}{C} = \frac{1}{C}\int_0^{t_1} I\,dt = \frac{I}{C}\int_0^{t_1} dt = \frac{I}{C}t_1 \tag{8.4}$$

   where $q$ is the charge of the capacitor.
2. Excitation of the LC circuit: After the short time $t_1$ the switch $SW_1$ will be switched off and the switch $SW_2$ will be switched on. The energy transfer from C to L starts up and the operating mode with steady state oscillations (no energy loss) is initiated. A charge/discharge current permanently flows between C and L. The expression of the voltage drop $v$ on the inductor:

$$v_L = -L\frac{di}{dt} \tag{8.5}$$

leads to the following differential equation of the charge $q$ of the capacitor:

$$L\ddot{q} + \frac{q}{C} = 0 \tag{8.6}$$

By writing

$$\omega_0^2 = \frac{1}{LC} \tag{8.7}$$

(8.6) is transformed into an expression identical to (8.2):

$$\ddot{q} + \omega_0^2 q = 0 \tag{8.8}$$

The general solution of the differential equation (8.8) has the well-known form:

$$q = A\cos\omega_0 t + B\sin\omega_0 t \tag{8.9}$$

where $A$ and $B$ are constants of integration determined by the initial conditions. If for $t = 0$, the electric charge and its derivative are: $q = q_0$ and $\dot{q} = \dot{q}_0$, then:

$$q = q_0\cos\omega_0 t + \frac{\dot{q}_0}{\omega_0}\sin\omega_0 t = K\cos(\omega_0 t + \alpha) \tag{8.10}$$

$$\dot{q} = \dot{q}_0\cos\omega_0 t - q_0\omega_0\sin\omega_0 t = -K\omega_0\sin(\omega_0 t + \alpha) \tag{8.11}$$

The parameters in the compact form are:

$$K = \sqrt{A^2 + B^2} = \sqrt{q_0^2 + \frac{\dot{q}_0^2}{\omega_0^2}} \text{ and } \alpha = \tan^{-1}\frac{-B}{A} = \tan^{-1}\left(-\frac{\dot{q}_0}{\omega_0 q_0}\right) \tag{8.12}$$

It is seen that the dependence of charge $q$ on time is the familiar sinusoid. Such a sinusoidal oscillation is characterized by three quantities:

1. Amplitude of the oscillations $K$;
2. Angular frequency $\omega_0$ (number of oscillations in $2\pi$ seconds);
3. Initial phase $\alpha$.

The amplitude and phase are determined by the initial conditions, but the angular frequency does not depend on these being determined by the parameters of the oscillatory system. An oscillatory state does not arise, when $q = 0$ and $\dot{q} = 0$, then the oscillator at the initial instant is in a state of equilibrium and remains so.

The formulas in (8.12) give an exact quantitative description of the LC harmonic oscillator defined by (8.6). They enable the user to calculate the values of $q$ and $\dot{q}$ for every instant of time if they are known at the instant of time $t = 0$.

A skillful parallel graphical analysis preferred by [3] is offered by the representation on the phase plane. To define a general application of this method, the quantities $q$ and $\dot{q}$ are renamed as follows:

$$x = q_0 \text{ and } y = \dot{q}_0 \tag{8.13}$$

The generalized variables $x$ and $y$ can be arbitrary time-dependent variables: a quantity $x$ and its velocity $y$ and are orthogonal Cartesian coordinates in the phase plane ($x$, $y$ plane of the states). The second name is given by the fact that to each state of the system, to, for example, each pair of values $x$ and $y$, there corresponds one and only one point on the plane. A variation of the system state corresponds to a motion of the corresponding representative point on the phase plan. The path followed by the representative point is called phase path and its velocity is called phase velocity. Finally, the curve described by representative points through the whole time (from $t = -\infty$ to $t = +\infty$ is called the complete phase path.

Having the solution of the differential equation of the harmonic LC oscillator (8.6), the equation of the phase path is easily found. The solution consists of two parametric equations (8.10) and (8.11), and the equation of the phase path results by eliminating the time $t$ from these equations:

$$\frac{x^2}{K^2} + \frac{y^2}{\omega_0^2 K^2} = 1 \tag{8.14}$$

If the parameter $K$ varies, this is the equation of a family of similar ellipses (i.e., with a constant ration of the axes) so that through each point of the phase plane there passes one and only one ellipse. The plane is filled with ellipses enclosing each other. Figure 8.7 is a specific representation of (8.14) for four values: $K = 1$ to 4, respectively.

Relative to the movement of the representative point along any phase path, this motion always occurs clockwise, since in the upper semiplane $y > 0$ and $x$ increases with time, while in the lower semiplane $y < 0$ and hence $x$ decreases with time.

The position-vector $\underline{\mathbf{V}}$ corresponding to an arbitrary representative point helps to define the phase velocity. Defining the position-vector as:

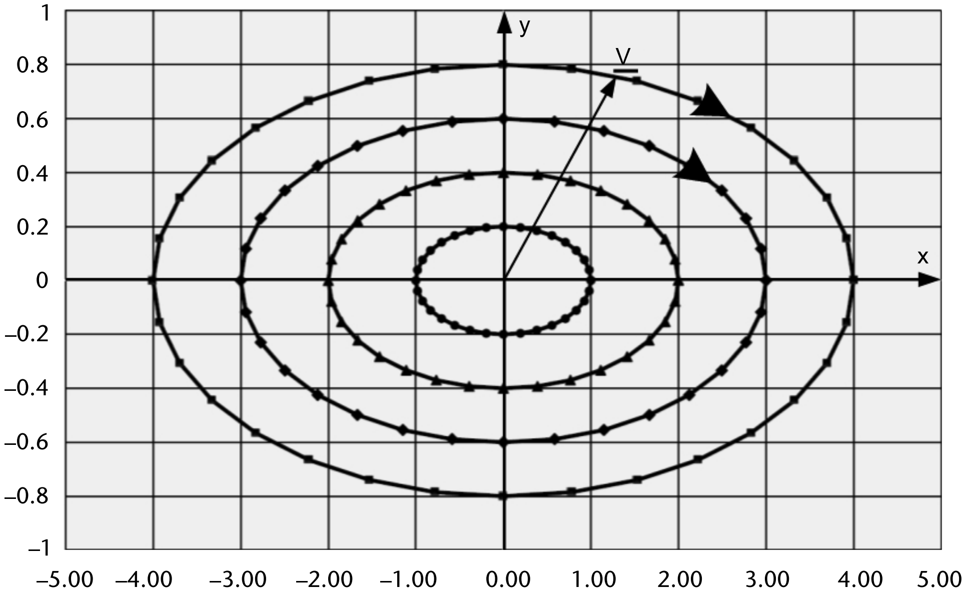


**Figure 8.7** Phase paths of the harmonic oscillator (square is at $K = 4$, diamond is at $K = 3$, triangle is at $K = 2$, and circle is at $K = 1$).

$$\underline{\mathbf{V}} = \underline{\mathbf{i}}x + \underline{\mathbf{j}}y \tag{8.15}$$

the velocity is represented in the form:

$$\underline{\mathbf{v}} = \frac{d\underline{\mathbf{V}}}{dt} = \underline{\mathbf{i}}\dot{x} + \underline{\mathbf{j}}\dot{y} \tag{8.16}$$

or according to (8.10) and (8.11), in the form:

$$\underline{\mathbf{v}} = \underline{\mathbf{i}}\left[-K\omega_0 \sin\left(\omega_0 t + \alpha\right)\right] + \underline{\mathbf{j}}\left[-K\omega_0^2 \cos\left(\omega_0 t + \alpha\right)\right] \tag{8.17}$$

With the exception of the case $K = 0$, the phase velocity never reduces to zero, since the sine and cosine never reduce to zero at the same time. The representative point accomplishes a whole circuit in $T_0 = 2\pi/\omega_0$ units of time. A phase path degenerating into a point corresponds to a state of oscillator equilibrium.

To conclude, the phase plane method provides the following benefits:

- Specific to graphical analysis methods, the mapped-out phase paths give an easily interpreted portrait of the system;
- The method can directly derive conclusions (e.g., from (8.2)) about the motion of the representative point without knowledge of the analytical expressions of the solutions of the initial equations (8.10) and (8.11).

This way, to directly arrive at the phase plane portrait from the initial system equation, it can also be illustrated for the harmonic oscillator. The quantity $y$ is already defined:

$$y \overset{\text{def}}{=} \frac{dx}{dt} \tag{8.18}$$

and the derivative of this formula associated with (8.2) leads to:

$$\frac{dy}{dt} = \frac{d^2x}{dt^2} = \ddot{x} = -\omega_0^2 x \tag{8.19}$$

Thus, the initial equation of the second order (8.2) is replaced by two equivalent equations of the first order. Dividing one of these equations by the other, the time dependence is eliminated and the resulting equation:

$$\frac{dy}{dx} = -\omega_0^2 \frac{x}{y} \tag{8.20}$$

defines the so-called integral curves, namely such curves that the tangents at every point of them have a slope (*dy*/*dx*) calculated from (8.20). On integrating (8.20), the formula of the phase path results, no longer in a differential form but in a finite form.

The method can be applied not only to a characteristic quantity and its derivative with respect to *t* but also to two interdependent system quantities.

Reloading the harmonic LC circuit (Figure 8.6(a)), a characteristic quantity for the capacity *C* is its charge *q*, whose derivative is the capacitor current $i_C$:

$$\frac{dq}{dt} = i_C \tag{8.21}$$

The magnetic flux is a characteristic quantity for the inductance *L* and can be expressed using the formula (3.7):

$$\Phi = \frac{V_m}{R_m} = \frac{n \cdot i_L}{R_m} = \frac{L}{n} i_L \tag{8.22}$$

where $V_m$ is the magnetomotive force (measuring unit: ampere-turns), *n* is number of turns, $i_L$ is the current through the inductor wire, and $R_m$ is the magnetic reluctance (see (3.2)).

The derivative of (8.22):

$$\frac{d\Phi}{dt} = \frac{L}{n} \cdot \frac{di_L}{dt} \tag{8.23}$$

defines the second state equation.

Obviously, the following equations are valid for the LC circuit in Figure 8.6(a), where:

$$i_L = i_C \text{ and } v_L = -v_C \tag{8.24}$$

and $v_L$ and $v_C$ are the voltage on the inductor and capacitor, respectively. These equalities help to eliminate time dependences in (8.21) and (8.23) to rewrite them and to highlight the state variable: $q$ and $\Phi$ as well as the parameters $L$ and $C$:

$$\dot{q} = n \cdot \frac{\Phi}{L} \text{ and } \dot{\Phi} = \frac{-1}{n} \cdot \frac{q}{C} \tag{8.25}$$

If the capacitance $C$ and the inductance $L$ are constant, the deterministic trajectories of the system on plane of the states with orthogonal Cartesian coordinate: $\Phi$ and $q$ are concentric circles (see Figure 8.7) having diameters dependent on $C$ and/or $L$ (important attribute for the inductive sensing elements). As long as the trajectories are convergent—a valid property for circles—the system is stable (solid-state oscillations).

### 8.2.2 Linear Oscillator with Losses

In opposition to the harmonic oscillator, which is characterized by a lossless resonant circuit, the real oscillators have lossy resonant circuits and consequently a different behavior. To examine them, [3] proposes to reconsider the mechanical example in Section 8.2.1 consisting of a body of mass $m$, which horizontally moves along a rod under the action of two lateral springs. The new assumption is that the motion is subject to friction caused by air and the frictional force is proportional to velocity.

The new equation of the motion will be:

$$m\ddot{x} + b\dot{x} + kx = 0 \tag{8.26}$$

where $b$ is the frictional coefficient.

This equation models the electrical analogue system, the Thomsonian circuit with ohmic loss resistance $R_S$ (see Table 8.1). This system obeys the equation (see (8.6)):

$$L\ddot{q} + R_S\dot{q} + \frac{q}{C} = 0 \tag{8.27}$$

where $q$ is again the charge on the capacitor and $L$, $R_S$, and $C$ are the inductance, resistance and capacitance, respectively.

Introducing the notations:

$$\frac{b}{m} = 2h,\ \frac{k}{m} = \omega_0^2 \text{ or correspondingly } \frac{R}{L} = 2h,\ \frac{1}{LC} = \omega_0^2 \tag{8.28}$$

the result is a general expression valid for (8.26) or (8.27):

$$\ddot{x} + 2h\dot{x} + \omega_0^2 x = 0 \tag{8.29}$$

The general solution of this differential equation has the well-known form:

$$x = Ae^{\lambda_1 t} + Be^{\lambda_2 t} \tag{8.30}$$

where $\lambda_1$ and $\lambda_2$ are the roots of the algebraic quadratic characteristic equation:

$$\lambda^2 + 2h\dot{\lambda} + \omega_0^2 = 0 \tag{8.31}$$

Depending on the sign of the equation determinant $\{h^2 - \omega_0^2\}$, there are two types of solutions and three different processes, which are

1. A damped oscillatory process, if the determinant is negative ($h^2 < \omega_0^2$);
2. An undamped oscillatory process, if the determinant is zero ($h^2 = \omega_0^2$);
3. A damped aperiodic process, if the determinant is positive ($h^2 > \omega_0^2$).

The study of the damped oscillatory process is important for electronic oscillators. For the condition $h^2 < \omega_0^2$ the above-mentioned roots are:

$$\lambda_{1,2} = -h \pm j\left(h^2 - \omega_0^2\right)^{1/2} = -h \pm j\omega \tag{8.32}$$

and the general solution of (8.29) is now (see (8.30)):

$$x = e^{-ht}\left(A\cos\omega t + B\sin\omega t\right) \tag{8.33}$$

where $A$ and $B$ are determined by the initial conditions (e.g., at $t = 0$, $x = x_0$, $\dot{x} = \dot{x}_0$):

$$A = x_0 \text{ and } B = \frac{\dot{x}_0 + hx_0}{\omega} \tag{8.34}$$

Similar to (8.10) and (8.11) the expressions of $x$ (8.33) and $\dot{x}$ can be written in a compact form:

$$x = Ke^{-ht}\cos(\omega t + \alpha) \tag{8.35}$$

$$\dot{x} = -K\omega_0 e^{-ht}\sin(\omega t + \alpha + \vartheta) \tag{8.36}$$

where the parameters are:

$$K = \sqrt{A^2 + B^2} \text{ and } \alpha = \tan^{-1}\left(\frac{-B}{A}\right) \tag{8.37}$$

and the quantity $\vartheta$ is determined by the relations:

$$h = \omega_0 \cdot \sin\vartheta \text{ and } \omega = \omega_0 \cdot \cos\vartheta \tag{8.38}$$

Formulas (8.35) and (8.36) define a damped oscillatory process in which

- The damping of the amplitude obeys an exponential law. The time constant of the exponential function depends on the damping coefficient $h$:

$$\frac{1}{h} = \frac{2L}{R_S} \tag{8.39}$$

- The damped oscillation has the conditional period:

$$T = \frac{2\pi}{\omega} = \frac{2\pi}{\left(\dfrac{1}{LC} - \dfrac{R_S^2}{4 \cdot L^2}\right)^{1/2}} \tag{8.40}$$

It should be noted that the measuring method described in Figure 4.12(c, d) applies this damped oscillatory operation mode.

The phase plane portrait of this system results from representing the formulas (8.35) and (8.36) in the phase plane. The result is a family of spirals having an asymptotic point at the origin of the coordinates (Figure 8.8).

### 8.2.3 Oscillators with Loss Cancellation by Positive Feedback Operation

Applying a driving input to the electronic oscillator with losses overcomes the losses of the circuit and causes the resonant circuit to oscillate indefinitely.

There are more possibilities to do that but all can be described by a structure consisting essentially of two parts: (1) the frequency-determining section, which may be a resonant circuit with lumped inductances and capacitances (for inductive sensors this belongs to the inductive sensing element), and (2) the maintaining section, which supplies energy to the resonant circuit to keep it in oscillation. Energy oscillates between the inductor, which stores the energy as a magnetic field, and the capacitor, which stores it as an electrical field.

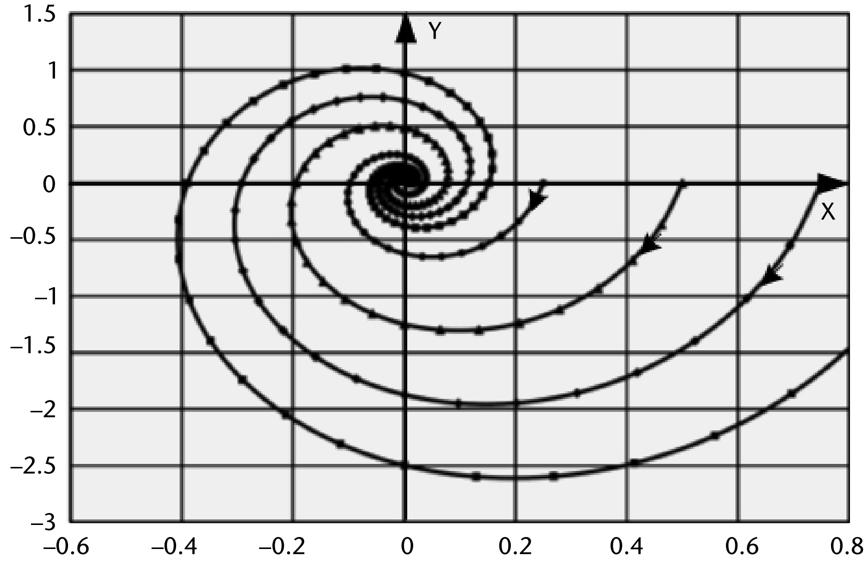


**Figure 8.8** Phase paths of the damped oscillator (square is at $K = 4$, diamond is at $K = 3$, triangle is at $K = 2$, and circle is at $K = 1$).

As previously seen, the impedance of the parallel resonant circuit is at resonance at a maximum and a small current should be driven from the maintaining section. This requires a DC source of power and consists of an electronic active section.

Certainly, the most popular method to get steady-state oscillations is to insert the resonant circuit with losses in a positive feedback loop of an electronic circuit. This method is presented in the current section. Other procedures like inclusion of the lossy oscillator in an AC bridge or relieve the losses connecting in parallel a negative resistor are subjects of the following sections.

To get a logical flow in the presentation of the positive feedback topic, the next step is to progress from the phase plane portrait toward the signal-flow graphs method [4] appertaining to the system/network theory.

Despite a graphical tool, this method permits also to gradually determine the analytical expression of a network transfer function making a step-by-step elimination of the formed graphic loops. The procedure is easier than classical analytical procedures. In that respect, the signal-flow graph is a diagrammatic representation of a linear system of equations and consists of two elements:

1. Node, which indicates a variable (node variable).
2. Branch, which indicates a relationship between a pair of variables $X_1$ and $X_2$, assigned to two nodes. The function relationship is called transmittance $G$ and the flow direction is given by an arrow on the branch. Supposing the arrow has a direction from the node $X_1$ to the node $X_2$, the transmittance is defined by:

$$G = \frac{X_2}{X_1} \tag{8.41}$$

The method has a reduced number of rules, and the most important to solve a graph are

- The value of a node variable is defined only by the incident branches entering it;
- Each incident branch contributes to the node value an amount equal to the branch transmittance multiplied by value of node variable from which the branch leaves;
- Sink nodes have only incident branches; source nodes have only branches that leave;
- In our case, for electrical networks, the node variables are generally voltages and/or currents and the transmittance are impedance, admittance, or their components.

The main elementary graph reductions are summarized in Table 8.2.

The positive feedback loop of an electronic circuit can be optimally defined starting with the representation of an ideal feedback amplifier as shown in Figure 8.9. The schematic has a general availability and consists of

- A basic amplifier, which may be any of four types: voltage, current, transconductance, or transimpedance amplifier and has a transfer function $A$;

- A feedback network with its transfer function $\beta$;
- The summing network (a real comparator or mixer) providing the comparison signal $X_d$ with the expression:

$$X_d = X_s \mp X_f \tag{8.42}$$

where the sign is minus if feedback signal $X_f$ is 180° out of phase with the source signal $X_s$ or plus if $X_f$ is in phase with $X_s$.

With the notation in Figure 8.9, the gain of the amplifier is:

$$A \equiv \frac{X_o}{X_i} \tag{8.43}$$

and the reverse transmission of the feedback network is defined by the expression:

$$\beta \equiv \frac{X_f}{X_o} \tag{8.44}$$

The gain with feedback of the system, obtained by a graph reduction in its signal-flow graph (Figure 8.10(a)) results in:

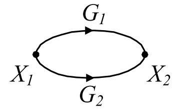
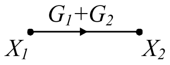
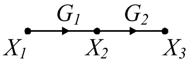
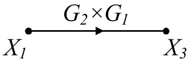
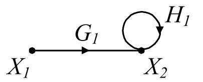
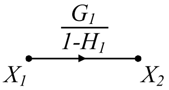
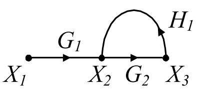
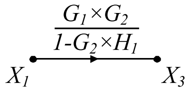
$$A_F \equiv \frac{X_o}{X_s} = \frac{A}{1 - \beta A} \tag{8.45}$$

When $\beta = 0$ (no feedback exists), (8.45) reduces to (8.43), then $A$ is frequently referred to as the open-loop gain and designated by $A_{OL}$.

When $\beta \neq 0$, the feedback loop exists and $A_F$ is often called closed-loop gain.

Similarly, the product $-\beta A$ is called loop gain or return ratio $T$:

**Table 8.2** Signal-Flow Graph: Elementary Graph Reductions

| *Graph Reduction* | *Initial Graph Segment* | *Reduced Graph Segment* |
|---|---|---|
| Parallel Configuration |  |  |
| Cascade Structure |  |  |
| Self-loop Configuration |  |  |
| Feedback Structure |  |  |

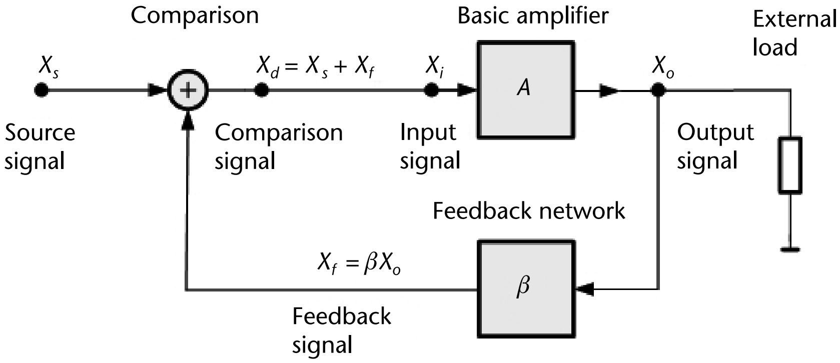

Comparison
Basic amplifier
External load
$X_s$
$X_d = X_s + X_f$
$X_i$
$A$
$X_o$
Source signal
Comparison signal
Input signal
Output signal
Feedback network
$X_f = \beta X_o$
$\beta$
Feedback signal

**Figure 8.9** General block diagram of the single-loop feedback amplifier model.

$$T \equiv -\beta A \tag{8.46}$$

Decisive is now the ratio between $A_F$ and $A$:

$$\text{If } |A_F| < |A| \text{ since } |1 - \beta A| > 1 \tag{8.47}$$

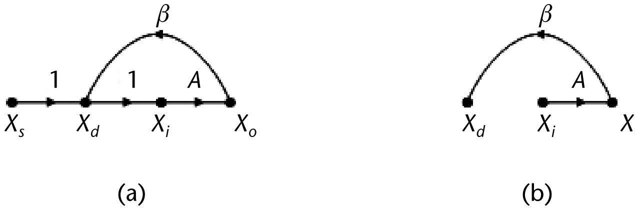
the feedback is termed negative. The extremely large volume of electronic amplifiers is based on this negative feedback. It gives system stability and provides many improvements (larger operating frequency range, reduction of noise and distortions, etc.).

$$\text{If } |A_F| > |A| \text{ since } |1 - \beta A| < 1 \tag{8.48}$$

the feedback is termed positive (regenerative). The operating mode of the majority of electronic oscillators implies this feedback type.

Three fundamental conditions are implicit in the representation of the ideal feedback amplifier (Figure 8.9):

1. The feedback network is unilateral, which means the source input signal is transmitted to the output through the amplifier and not through the feedback network;
2. The amplifier is also unilateral and transmits only from its input to output, and thus the feedback signal is transmitted only through the feedback network.


$\beta$
1
1
$A$
$X_s$
$X_d$
$X_i$
$X_o$
$\beta$
$A$
$X_d$
$X_i$
$X_o$

(a) (b)

**Figure 8.10** (a) Signal-flow graph constructed for the single-loop feedback amplifier model, and (b) signal-flow graph for open-loop amplifier and zero input signal *Xs*.

3. The transfer ratio $\beta$ is independent of the source and load resistances $R_S$ and $R_L$.

Practical feedback circuitry only approximates these conditions. The feedback networks usually consist of passive elements and thus are not absolutely unilateral. Both units $A$ and $\beta$ are generally affected by the load and source impedances.

The next fundamental question is at which frequency a sinusoidal oscillator will operate if the condition (8.48) is fulfilled. To find the expression of this oscillation frequency, it is necessary to develop the expressions of $A$, $\beta$, $T$, and so forth, taking into consideration their frequency dependence.

These values are generally complex quantities and to prevent the computation with the imaginary variable $j$ it is advisable to replace the angle frequency $\omega = 2\pi f$ by the Laplace operator [4] defined by the formula:

$$s \equiv \sigma + j\omega \tag{8.49}$$

where $\sigma$ is the Neper frequency in nepers per second and $\omega$ is the angular frequency in radians per second [5]. In steady-state $\sigma = 0$ and therefore $s = j\omega$.

The benefit of this variable and of the work in the $s$-domain is the transforming of the equations with complex quantities in algebraic equations. Steady-state impedance for an inductor having inductance $L$ or for a capacitor with capacitance $C$ is $\underline{Z} = sL$ and $\underline{Z} = 1/sC$, respectively (see (4.4)).

Thus, the differential equations, which describe the system behavior versus frequency, become algebraic equations and can easily be solved.

Finally, getting the algebraic expression of the transfer function, the return from the $s$-domain to the time domain, to determine the magnitude and phase of the transfer function is possible by applying an inverse Laplace transform [4].

As an example of the convenience of this method, let us suppose the single loop feedback amplifier in Figure 8.9 consists of an ideal basic amplifier without frequency dependence and a feedback network with a second-order frequency dependence with the corner frequencies $\omega_1$ and $\omega_2$.

The frequency dependence of the return ratio results in:

$$T(j\omega) = -\frac{A_0 \cdot \beta_0}{\left(1 + \dfrac{j\omega}{\omega_1}\right) \cdot \left(1 + \dfrac{j\omega}{\omega_2}\right)} = -\frac{A_0 \cdot \beta_0 \cdot \omega_1 \cdot \omega_2}{\left(\omega_1 + j\omega\right) \cdot \left(\omega_2 + j\omega\right)} \tag{8.50}$$

where $A_0$ and $\beta_0$ are the gain of the amplifier and the reverse transmission of the feedback network at very low frequencies, respectively.

The substitution of (8.49) with $\sigma = 0$ into (8.50) gives the algebraic equation:

$$T(s) = -\frac{A_0 \cdot \beta_0 \cdot \omega_1 \cdot \omega_2}{\left(\omega_1 + s\right) \cdot \left(\omega_2 + s\right)} = -\frac{T_0}{s^2 + s \cdot \left(\omega_1 + \omega_2\right) + \omega_1 \cdot \omega_2} \tag{8.51}$$

Substituting now (8.51) into (8.45) and making reductions, the closed-loop gain with feedback $A_F$ in terms of $s$ results in:

$$A_F(s) = \frac{N(s)}{D(s)} = \frac{N(s)}{s^2 + s \cdot (\omega_1 + \omega_2) + \omega_1 \cdot \omega_2 + T_0} \tag{8.52}$$

The roots of the numerator $N(s)$ are called zeros and the roots of the denominator $D(s)$ are called poles of the system (the points of infinite response). They are of great importance for systems with feedback, since they clarify a number of aspects including the characteristic frequencies. The distribution of the poles into the $s$ plane gives a view about the system features.

Equation (8.52) indicates that the poles of $A_F(s)$ are function of $T_0$, the amount of feedback. Figure 8.11 shows the movement of the closed-loop poles as $T_0$ is increased. They start at $-\omega_1$ and $-\omega_2$, the poles of $T(s)$, and move toward each other along the negative real axis as $T_0$ is increased from zero. The poles coincide at the frequency $-(\omega_1 + \omega_2)/2$. Thus, $A_F$ has two equal poles.

If $T_0$ is further increased, the poles become complex with the real part remaining at $-(\omega_1 + \omega_2)/2$. The closed-loop poles always remain in the left half-plane, which is a system stability indicator.

Considering now a second example a with a positive feedback, if the closed-loop poles result into the right half-plane, the feedback amplifier becomes unstable; that is, it oscillates. No excitation is needed to produce an output. If the closed-loop poles can be made to lie on the $j$-axis, the natural response of the system is a sinusoid whose frequency is the pole frequency. This idea is the basis of sinusoidal oscillator circuits.

In other words, a sinusoidal oscillator is a feedback amplifier designed to have closed-loop poles on the $j$-axis at the need frequency of the output. Once the poles tend to move into the right half-plane ($T_0$ still increases), their real part forces an increasing of the exponential output buildup. However, the voltage is bounded by limitations (intrinsic or implemented). At the peaks of the sine wave, the limitation reduces the loop gain, driving the system poles from the right into the left half-plane. The amplitude decays, the circuits reenter the active region, and roots move back into the right half-plane. The oscillator will be self-starting only if the poles initially lie in the right half-plane.

To illustrate the concept of an oscillator based on feedback amplifier, the amplifier graph in Figure 8.10(a) can be changed opening the loop and considering a source signal equal to zero (Figure 8.10(b)). Now a signal $X_i$ is applied directly into

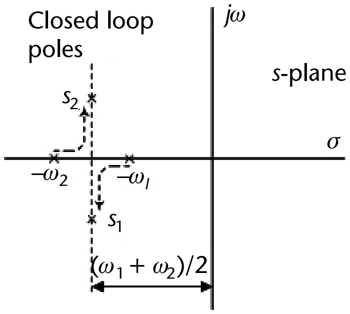


**Figure 8.11** Root locus of the two-pole transfer function (8.52) and pole migration as a function of $T_0$.

the amplifier and amplified in this stage as well as in the feedback network. The output signal of the summing network is now:

$$X_d = \beta X_o = \beta A X_i = T X_i \tag{8.53}$$

where $T$ is the return ration of the system.

Suppose it should happen that matters are adjusted in such way the signal $X_d$ is identical to the externally applied signal $X_i$. Since the basic amplifier has no means to distinguishing the source of the input signal applied to it, it would appear that if the external source were removed and if the node $X_d$ were connected to the node $X_i$, the basic amplifier would continue to provide the same output signal $X_o$.

The statement "identical" means that the instantaneous values of both signals are exactly equal at all times. This condition is equivalent to the stipulation that the return ratio must be equal to one at all times:

$$T(s) = -1 \tag{8.54}$$

For a sinusoidal output waveform, the condition (8.54) is equivalent to the condition that the amplitude, phase, and frequency of $X_d$ and $X_i$ must be identical.

Hence, a second answer to the question about the value of the oscillator working frequency results: The frequency at which a sinusoidal oscillator will operate is the frequency $f_0$ for which the complex return ratio fulfills the condition:

$$\underline{T}(j2\pi f_0) = -1 \tag{8.55}$$

Alternatively, this condition can be expressed using the Cartesian terms:

$$\Re\left\{\underline{T}(j2\pi f_0)\right\} = -1 \text{ and } \Im\left\{\underline{T}(j2\pi f_0)\right\} = 0 \tag{8.56}$$

or using the polar representation, which is much more popular:

$$\left|\underline{T}(j2\pi f_0)\right| = 1 \text{ and } \angle\underline{T}(j2\pi f_0) = -180^\circ \tag{8.57}$$

The condition (8.55) is called the Barkhausen criterion and its development in polar components (8.57) states that two conditions must be satisfied by a circuit to sustain oscillations:

1. The magnitude of the gain of the basic amplifier and feedback network must be equal to one.
2. The phase shift through the basic amplifier and the feedback network must be 360° (or $2\pi n$ radians, where $n$ is a natural number $n = 1, 2, \ldots$). Recall that the definition of $T$ includes a minus sign that is equivalent to 180° of phase shift.
3. To these theoretical conditions a third general principle dictated by practical considerations should be mentioned. For $|\underline{T}|$ value smaller than unity, the

removal of the external generator will result in a cessation of oscillations. For larger values, the output signal would theoretically increase after every rotation through loop. The amplitude of oscillations is somewhat limited through the nonlinearity of the electronic components, temperature deviations, and so forth. In conclusion, in a practical oscillator the loop gain should be adjusted slightly larger than unity (to prevent a falling drop below unity), and the amplitude of the oscillations is limited by the onset of nonlinearity.

At the end, it should underline the two applications of the Barkhausen criterion: (1) To prevent classical amplifiers from breaking out into oscillation, and (2) to determine how to cause the system to oscillate and how to find the frequency of the sinusoid generated for no excitation applied to the feedback system. This application is relevant to our oscillator's topic.

## 8.3 Convenient Types of LC Oscillators for Inductive Sensors

In the technical literature there is plenty of information on electronic oscillators and some attempts to classify them. An important category, sinusoidal oscillators, has widespread use in electronics and sensor technique.

The following author's original classification is made from the point of view of the oscillators' utilization in inductive sensors and therefore is limited with regard to the field of oscillators with LC reactive components.

The major classification criterion refers to the type of frequency-determining section and/or of the maintaining section (see definitions at the beginning of Section 8.2.3). Accordingly, the following main classes and subclasses can be specified as follows:

a. Oscillators having an amplifier and positive feedback network containing
    - LC quadripoles (e.g., Hartley-, Colpitts-, Clapp-oscillator, etc.)
b. Oscillators with feedback LC dipoles supplied by AC-currents:
    - LC dipole in ring-circuit oscillators
    - LC dipole in oscillators with differential amplifier;
    - LC dipole in bridge oscillator architecture
    - LC dipole with relaxation oscillators
c. Oscillators with LC-dipoles with pulsing DC-current excitation
d. Oscillators with LC dipoles and loss compensation by negative resistance:
    - LC dipole connected to a device having a negative region of its volt-ampere characteristic
    - LC dipole losses relieved by circuitry with additional negative resistance emulation
e. Miscellaneous individual company proprietary versions.

The first three classes (a. to c.) can be joined together as oscillators with positive feedback (Sections 8.3.1 to 8.3.5), while the fourth class (d.) flags the oscillators with negative resistance (Section 8.3.6).

A second criterion is absolutely specific for the application of the oscillators in inductive sensors. It is barely known and refers to the oscillator behavior if the target moves inside the sensing range. Two oscillator types can be distinguished:

1. Pulling-down oscillators, which have an output signal-distance characteristic with an unwanted cutoff point. Below or above this specific distance the oscillator output signal becomes zero. This behavior is generally undesired and exclusively restricts the oscillator application area to inductive proximity switches (Section 2.1.1).
2. Oscillators with a flat signal-distance characteristic, which perform steady-state oscillations over the entire sensing range. Obviously, they are suitable both for analog and binary sensor realizations.

The requested features of various oscillators to make them suitable for inductive sensor applications can be summarized as follows:

- High sensitivity (see Section 1.3.3) to perform stable target detection over large sensing ranges;
- Greater—and if possible linear—adjustability to provide reproducible quality of the sensor alignment;
- Schematics suitable for integration (bipolar and/or CMOS; see Sections 7.3.1 and 7.3.2) to facilitate sensor miniaturizations;
- High immunity against harsh environment influences specific to industrial applications (see Chapter 2);
- Low current consumption and compliance with the restrictive supply conditions stated by the standards (ranges, unipolar supply voltage, etc.);
- High signal-to-noise ratio;
- Cost saving and high reliability as a result of the low number of components and connections.

### 8.3.1 Transistor-Based, Positive-Feedback Oscillators

Traditional transistor-based oscillators with feedback LC quadripoles, also called three connection-point oscillators, were long the preferred solution for inductive sensors with discrete components (see Section 7.2). Today they have more of a historical relevance, as they are neither better-performing nor more improved.

However, to keep to the initial goal of providing an exhaustive and evaluative study, the chapter starts again in this section with a summary of the transistorized discrete version in a similar matter as in Chapter 7 (see Section 7.2).

In addition, as we know, this technology is still currently in use on the low-cost market segment, after all.

These numerous oscillator types will be initially introduced by means of a general equivalent circuitry and the corresponding equations, which should be valid for all types. Minor specific features of every type will be stated after that.

Considering an amplifying stage (at the limit, with a simple bipolar transistor), the amplifier can be converted into an oscillator inserting a positive feedback network $\underline{Z}_1$ to $\underline{Z}_3$ as shown in Figure 8.12(a). The topologies of this passive quadripole are very different.

The amplifier in Figure 8.12(a) is represented in Figure 2.18(b) by an ideal op-amp having a very high input impedance (infinite), a finite output resistance $R_O$, and a linear high gain, which is not yet affected by the frequency at oscillator resonance frequency (its 3-dB cutoff frequency is significantly higher).

The components of the return ratio $T$ defined in (8.46) are now the gain of the amplifier (see (8.43)):

$$\underline{A} = \frac{\underline{v}_2}{\underline{v}_1} \tag{8.58}$$

and the reverse transmission of the feedback network $\beta$ (see (8.44)) that can now be expressed in terms of its specific topology in Figure 8.12(b):

$$\underline{\beta} = \frac{\underline{v}_2^*}{\underline{v}_2} = \frac{\underline{Z}_1\underline{Z}_2}{R_O\left(\underline{Z}_1 + \underline{Z}_2 + \underline{Z}_3\right) + \underline{Z}_1\left(\underline{Z}_2 + \underline{Z}_3\right)} \tag{8.59}$$

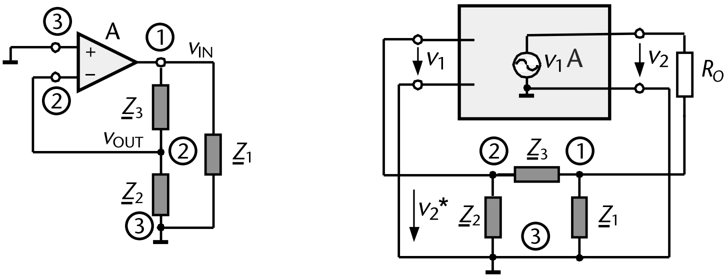
According to Barkhausen criterion, the circuit indefinitely oscillates if at the resulting resonance frequency $f_0$ the return ratio $T$ is real and is equal to minus one (see (8.55)):

$$\underline{T}\left(j2\pi f_0\right) = \underline{A}\left(j2\pi f_0\right) \cdot \underline{\beta}\left(j2\pi f_0\right) = -1 \tag{8.60}$$

With (8.59), (8.60) gets its specific expression:

$$\underline{A} \cdot \frac{\underline{Z}_1\underline{Z}_2}{R_O\left(\underline{Z}_1 + \underline{Z}_2 + \underline{Z}_3\right) + \underline{Z}_1\left(\underline{Z}_2 + \underline{Z}_3\right)} = -1 \tag{8.61}$$

It is important to note that the minus sign in the right equation part reflects the 180° phase rotation caused by the used inverting input of the amplifier.



(a) (b)

**Figure 8.12** (a) Principle block diagram of an electronic oscillator with LC feedback quadripole, and (b) equivalent circuit.

In the present case, the impedances have only reactive parts and can generally be expressed by:

$$\underline{Z}_i = jX_i \quad (i = 1,2,3) \tag{8.62}$$

where $X_i$ are either the corresponding inductive reactances ($\omega L$) or the capacitive reactances ($-1/\omega C$).

The substitution of (8.62) into (8.61) gives:

$$A \cdot \frac{-X_1X_2}{jR_O(X_1 + X_2 + X_3) - X_1(X_2 + X_3)} = -1 \tag{8.63}$$

From (8.63) results the condition for a real return ratio:

$$X_1 + X_2 + X_3 = 0 \tag{8.64}$$

Under this condition, the formula (8.63) becomes:

$$A \cdot \frac{X_2}{(X_2 + X_3)} = -1 \tag{8.65}$$

Making a final substitution of (8.64) into (8.65), this expression gets its definitive form:

$$A \cdot \frac{X_2}{X_1} = 1 \tag{8.66}$$

The formula (8.64) can be used to determine the resonance frequency $f_0$ and (8.66) represents the design condition to obtain steady-state oscillations.

As mentioned above, this analysis is based on ideal models for the maintaining section as well as for the frequency-determining section (lossless lumped inductances and capacitances). However, the equations provide a sufficient accuracy for the praxis. Higher exactitude is offered by the CAE tools like SPICE simulation software, as will be exemplified in the following sections.

If $X_1$ and $X_2$ are inductances and $X_3$ is a capacitance, the circuit is called a Hartley oscillator (Figure 8.13(a)).

With the component names in Figure 8.13(a), the reactances $X_1$ to $X_3$ become:

$$X_1 = j\omega L_1;\ X_2 = j\omega L_2;\ X_3 = j\frac{-1}{\omega C} \tag{8.67}$$

Making a final substitution of (8.67) into (8.64), the resulting equation:

$$j\omega_0 L_1 + j\omega_0 L_2 + j\frac{-1}{\omega_0 C} = 0 \tag{8.68}$$

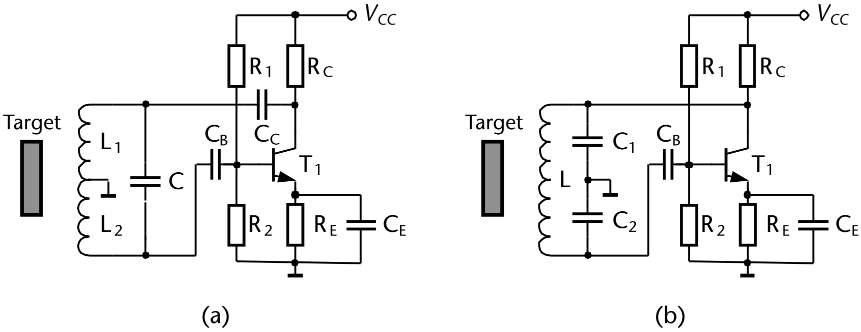


**Figure 8.13** The most popular oscillators with bipolar transistors in common-emitter configuration used in inductive sensors: (a) Hartley version, and (b) Colpitts version.

allows to immediately determine the expression of the resonance frequency $f_0$ of the Hartley oscillator:

$$f_0 = \frac{1}{2\pi}\sqrt{\frac{1}{\left(L_1 + L_2\right)C}} \tag{8.69}$$

Finally, the formula (8.66) allows to determine the design rule of the Hartley oscillator:

$$A\left(f_0\right) = \frac{L_2}{L_1} \tag{8.70}$$

The oscillator circuitry in Figure 8.13(a) shows besides the above-mentioned Hartley feedback network, $L_1$, $L_2$, and C, the classical biasing network of the bipolar transistor $T_1$ used in common-emitter configuration (resistors $R_1$, $R_2$, $R_C$, and $R_E$), the coupling capacitors $C_B$ and $C_C$, and the bypass capacitor $C_E$ for emitter bias resistor $R_E$.

For earthed tap requirement, the tap can be taken to GND, to the power supply, which has AC low impedance do to its always-used bypass capacitor, or any point with a low AC impedance. The ends of the inductor present voltages, which are in antiphase with each other. This is a Barkhausen's demand because the single transistor T1 is an inverting amplifier. Its gain should fulfill (8.70).

The circuit can thus be regarded as that of an amplifier, which supplies its own input; in other words, there is considerable positive feedback and this causes the signal building up rapidly. The output is taken from the transistor itself (e.g., from a resistor in the collector circuit).

If $\underline{Z}_1$ and $\underline{Z}_2$ in Figure 8.12 are capacitances and $\underline{Z}_3$ is an inductance, the circuit is called a Colpitts oscillator (Figure 8.13(b)). The reactances $X_1$ to $X_3$ are now:

$$X_1 = j\frac{-1}{\omega C_1};\ X_2 = j\frac{-1}{\omega C_2};\ X_3 = j\omega L \tag{8.71}$$

Following a similar procedure like for the Hartley oscillator, the final expressions for the resonance frequency $f_0$ and design rule of the Colpitts oscillator are:

$$f_0 = \frac{1}{2\pi}\sqrt{\frac{1}{L \cdot \frac{C_1 + C_2}{C_1 \cdot C_2}}} \tag{8.72}$$

and

$$A(f_0) = \frac{C_1}{C_2} \tag{8.73}$$

The circuitry in Figure 8.13(b) is very similar to circuitry in Figure 8.13(a), described above and does not need additional information. The capacitive branch of the feedback network has the central point connected to GND. The voltage *BE* is again in antiphase with the voltage *CE*.

The existence of coil tap represents a practical disadvantage of the Hartley oscillator in comparison with the Colpitts oscillator. The need of the tap increases the manufacturing costs. In addition, a high reproducibility of the obtained taps in the coil manufacturing is a challenging target for the high-volume production.

Considering an amplifying stage with the simplest bipolar-transistor common-emitter configuration, the amplifier can be converted into an oscillator inserting a positive feedback network, namely a feedback LC quadripole. The variety of transistor-based oscillators with feedback LC quadripoles is very large [2, 6].

An effective procedure to summarize this oscillator's field can be started beginning with a systematic résumé of the possible topologies of the three-component feedback LC networks (Figure 8.14). They are single C-topologies consisting of a

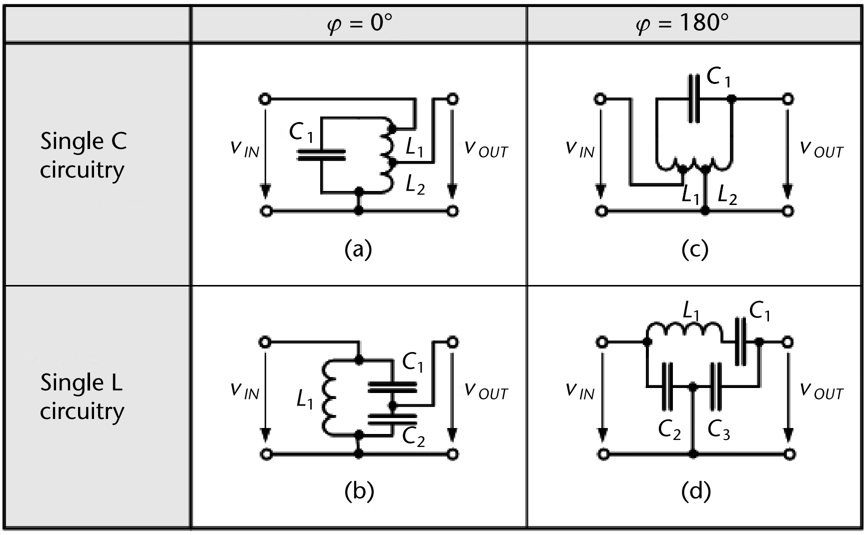


**Figure 8.14** Feedback LC quadripoles for the classical oscillators: (a) Hartley version, (b) Colpitts, (c) Lampkin, and (d) Clapp.

capacitor and a tapped inductor as well as single L topologies containing an inductor and a capacitor of two parts. The phase rotation value depends on the chosen third-connection point. The classical schematics carry traditional names.

Beside common-emitter amplifiers, common-base or common-collector amplifiers can be used. The amplifiers can also operate in class C for large oscillation amplitudes [2] or can use unipolar JFETs instead of bipolar transistors.

As an example, Figure 8.15 shows a Colpitts oscillator with bipolar transistor in common-base configuration. Since the output and the input voltages are now in phase with each other, there is no need for a phase rotation in the feedback quadripole (see Figure 8.14(b)). In a simple manner, the transistor is supplied through the feedback network.

Instead of single electronic component for each of the reactances $X_1$, $X_2$, and $X_3$ (Figure 8.12), some popular oscillators use LC dipoles with tuned behaviors. Primarily these dipoles are themselves parallel resonant circuits. Usually, the resonance frequencies of $X_1$ and $X_2$ are roughly the same and different from the resonance frequencies of $X_3$. This ensures that the behavior of $X_1$ and $X_2$ (capacitive or inductive) at the oscillator resonance is opposite to the behavior of $X_3$.

According to [2], the Meissner oscillator is the oldest oscillator circuit (1913) from the pioneering years of the radio engineering. The frequency-determining section consists of a parallel resonant LC circuit, which is directly connected in the collector circuit of a classical bipolar transistor in common-emitter configuration (similar to Figure 8.13(a) but without $R_C$). The inductance of this circuit is, practically speaking, the primary winding of a classical high-frequency transformer. The secondary winding is connected, via a coupling capacitor, in parallel to the BE circuit and provides the oscillator feedback. Fulfillment of the Barkhausen conditions (8.57) is very easy to ensure: (1) the phase condition is fulfilled by the adequate choice of the polarity of the secondary winding connections, and (2) the gain condition is met by the suitable fixing of the transfer ratio of the transformer. In addition, the circuit can typically be realized with JFET transistors (tuned-gate oscillator) in which the roles are reversed: the secondary transformer winding is directly inserted in the drain circuit and the primary winding is connected together with the resonance capacitor in parallel to the high-impedance gate input of the JFET. If the transformer is replaced by an autotransformer, the schematic becomes the Harley version.

A disadvantage appears when the input impedance of the maintaining section is lower than the output impedance of the frequency-determining section. This output impedance is shunted down and the Q factor of the existing resonant circuit is

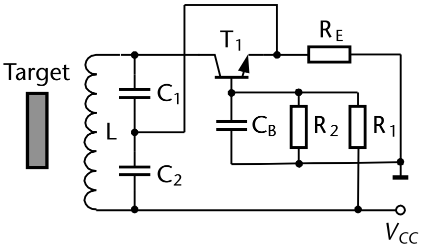


**Figure 8.15** Colpitts oscillator with bipolar transistor in common-base configuration.

thereby lowered. The Lampkin oscillator principle is based on a better impedance matching by using an inductor with two taps in the feedback quadripole (Figure 8.14(c)). The technique is to drive the input of the active device from a portion of the inductor that constitutes an approximate impedance match (impedance-matching tap). The second tap is the feedback tap.

In the Clapp oscillator (Figure 8.16), freedom from the shunting effect of the active device is achieved by using a series-resonant circuit, which inherently has a low impedance (see components $L_1$ and $C_1$ in Figure 8.14(d)). Apart from that, the schematic is very similar to the Colpitts oscillator. The resonance frequency $f_0$ depends on all capacitors $C_1$ to $C_3$ connected in series in the resonant circuit:

$$f_0 = \frac{1}{2\pi}\sqrt{\frac{1}{L \cdot C}} \tag{8.74}$$

where $C$ is expressed by:

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} \tag{8.75}$$

The Reinartz oscillator [6] belongs to the class with more complex feedback topology. In addition to a main passive parallel resonant circuit (not supplied, only grounded), the feedback network contains two inductors, which are coupled (transformer coupling) to this frequency-determining part and are connected in series with the collector and emitter, respectively. Positive feedback is achieved by coupling the collector circuit to the emitter circuit via mutual inductances between series inductors and the main inductor.

The technique of the Huth-Kuehn oscillator [2] can be considered related to the Reinartz technique. The oscillator contains two fully independent parallel LC resonant LC circuits connected to the BE input and to the CE output of the bipolar transistor in common-emitter configuration. Both circuits are dimensioned to have an inductive behavior (its own resonance frequency higher than the requested oscillator frequency; see Section 8.1.2). These subcircuits are equivalent to the reactances $X_1$ and $X_2$ (Figure 8.12). The reactance $X_3$ is realized by the transistor feedback capacitance $C_{bc}$, in a way contrary to the classical neutralization of an amplifier, which provides negative feedback over this path.

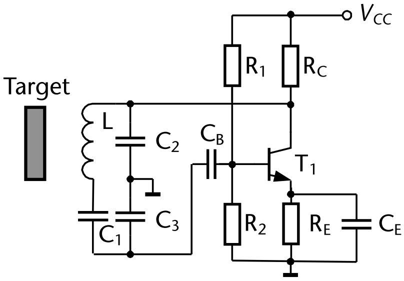


**Figure 8.16** Clapp oscillator with bipolar transistor in common-emitter configuration.

Looking at the topology of the feedback LC quadripoles in Figure 8.14, their Π topology is visible. Applying a triangle-star transformation, these schematics can be converted into *T* elements. Reference [2] describes oscillators with such feedback circuitries and details their considerably small significance for practical applications.

Before ending this section, it is useful to highlight [7], which contains—in the author's opinion—the more extensive knowledge of the field of these oscillator types. The reader can find here more information and details about this item.

### 8.3.2 Ring-Circuit Oscillators with LC Dipole

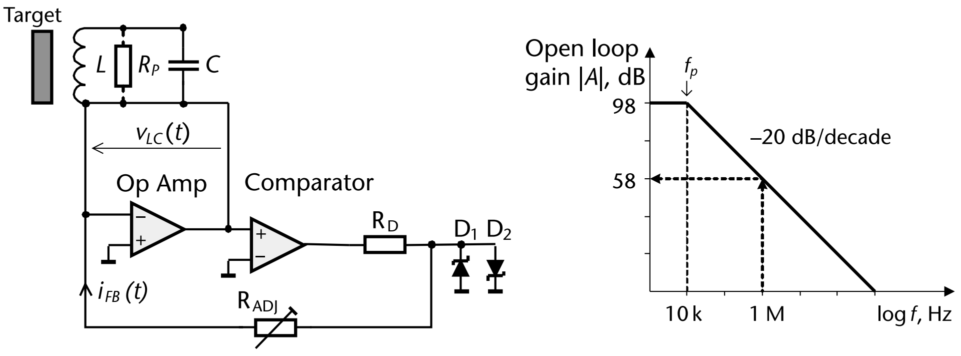
The schematic in Figure 8.17(a) represents a functional block diagram of this oscillator class with LC dipole as well as a specific example, which has numerous applications.

The oscillator feedback part is a classical parallel resonant LC circuit (see Section 8.1.2). The virtual resistor $R_P$ (represented by dashed line) symbolizes the parallel loss resistance (see (4.38) for $R_P$ definition and (5.35) for total $R_S$ expression). To provide a minimal impedance loading, this circuit is directly placed in the feedback loop of the op-amp. Considering the ideal op-amp with the specific assumptions: (1) the current to each input is zero and (2) the voltage between the two input terminals is zero, and having its noninverting terminal connected to GND, then the parallel resonant LC circuit rests on the virtual GND (inverting terminal) and freely swings with the amplitude $V_{lcm}$. Its excitation magnitude is an AC current provided by a voltage-controlled rectangular current source.

The op-amp output signal with the general expression:

$$v_{LC}(t) = V_{lcm} \sin\left(2\pi f_0 t + \varphi\right) \tag{8.76}$$

is led to a high-speed zero-cross comparator, which commutes between high and low levels of its output for every zero crossing at its noninverting input. For a better level control, Schottky barrier diodes (e.g., BAT42) clamp the levels on $\pm V_{SD} = 0.35\text{V}$ (typical voltage drop at forward biases in the range of a few milliamps [7]).



(a) (a)

**Figure 8.17** Ring-circuit oscillator with LC dipole: (a) circuitry, and (b) frequency response of the open-loop gain of the used commercial op-amp in a straight-line Bode approximation [7] provided by the op-amp manufacturer in the data sheet.

This symmetrical voltage rectangular wave with defined amplitude (±0.35V) independent of $V_{lcm}$ and in phase with $v_{LC}(t)$ is converted by means of the feedback adjustment resistor $R_{ADJ}$ into a symmetrical current wave:

$$i_{FB}(t) = \begin{cases} \dfrac{+V_{SD}}{R_{ADJ}} & \text{for } 0 < \left(\omega_0 t + \varphi\right) < \pi \\ \dfrac{-V_{SD}}{R_{ADJ}} & \text{for } \pi < \left(\omega_0 t + \varphi\right) < 2\pi \end{cases} \tag{8.77}$$

which is fed into the parallel resonant LC circuit.

The classical Fourier series representation [5] of the rectangular wave carries out to a sum of sinusoidal waves with frequencies $f_0$, $3f_0$, $5f_0$, and so forth:

$$i_{FB}(t) = \frac{4}{\pi} \cdot \frac{+V_{SD}}{R_{ADJ}} \left[ \sin\left(2\pi f_0 t + \varphi\right) + \frac{\sin 3 \cdot \left(2\pi f_0 t + \varphi\right)}{3} + \frac{\sin 5 \cdot \left(2\pi f_0 t + \varphi\right)}{5} \right] \tag{8.78}$$

The parallel resonant LC circuits for inductive sensors usually have high values of the $Q$ factor. Correspondingly, the bandwidth $B_Q$ (see Table 8.1) has low values (below a few kilohertz). Due to this high selectivity, the harmonics of the current wave are strongly suppressed. The only effective component is the fundamental wave, having the same phase and frequency $f_0$ as the $v_{LC}(t)$ and an amplitude according to the expression (8.78):

$$I_{fbm} = \frac{4}{\pi} \cdot \frac{+V_{SD}}{R_{ADJ}} \tag{8.79}$$

In summary, this current sets the amplitude of the oscillations and the final oscillations expression results:

$$v_{LC}(t) = R_P \cdot I_{fbm} \sin\left(2\pi f_0 t + \varphi\right) = \frac{4 \cdot V_{SD}}{\pi} \cdot \frac{R_P}{R_{ADJ}} \cdot \sin\left(2\pi f_0 t + \varphi\right) \tag{8.80}$$

*Example 8.1*

For the implementation of the oscillator shown in Figure 8.17(a), a good choice is the high-gain bandwidth, high-speed monolithic op-amp AD849 (manufacturer Analog Devices). Its characteristic open-loop gain versus frequency has a very high gain at low frequencies (≈ 100 dB) and a single dominant pole at $f_p = 10$ kHz (Figure 8.17(b)). That provides proper ideal op-amp conditions even for high frequencies (e.g., a gain of 58 dB at 1 MHz). The ultrafast precision comparator LT1016 (manufacturer Linear Technology Corporation) is also a suitable solution for a high-speed zero-cross detector in frequency ranges below a few of megahertz. The current through the Schottky diodes is limited by $R_D = 270\Omega$.

The immediate use of the oscillator circuitry in Figure 8.17(a) in serial-fabrication inductive sensors is conditionally recommended; the components are expensive and the need for symmetrical supply voltages is not compliant with the standards.

More relevant is the circuitry as a pilot oscillator for quick-and-dirty laboratory measurements. For a known fixed $R_{ADJ}$ value, the dependence $R_P$ versus distance $s$ can be obtained by measuring the oscillation amplitude $V_{lcm}$ at these different distances and using the formula:

$$R_P = \frac{\pi \cdot R_{ADJ}}{4 \cdot V_{SD}} \cdot V_{lcm} \tag{8.81}$$

Figure 8.18 shows a concrete record ($RP_{\text{measured}}$) of the dependence of $R_P$ versus distance $s$ between the coil and a reference steel target for an exemplifying resonant LC circuit with L = 190 $\mu$H and C = 1 nF.

An in-depth look at the curve shows that it carries to a postulate, according to which the curve has an exponential aspect and could accurately be approximated by a calculated dependence $R_{P\_clc}$ versus $s$ that is expressed by:

$$R_{P\_clc} = R_{P\infty}\left[1 - e^{-(s-s_0)/SK}\right] \tag{8.82}$$

where $R_{P\infty}$ is the measured free-running value of $R_P$; that is, the $R_P$ value with removed target or for target at very large distances to the coil, $s_0$ is the virtual intersection with the $x$-axis (abscissa) and $SK$ is an exponential constant characteristic for the system.

The determination of $SK$ is a graphical operation and relies on the well-known property of the natural exponential function, which this achieves 99.3% from the free-running value after five constants. Thus the expression of $SK$ is:

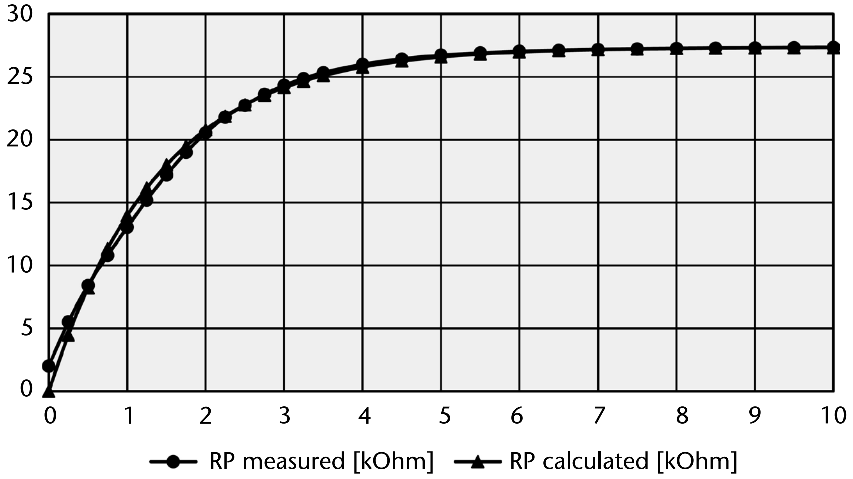


**Figure 8.18** Two plots of the characteristic $R_P$ versus distance $s$; $RP_{\text{measured}}$ (determined with the pilot oscillator in Figure 8.17(a) and (8.81)) as well as $RP_{\text{calculated}}$ (approximated with (8.82)) for $\{R_{P\infty} = 27.35\ \text{k}\Omega,\ SK = 1.4\ \text{mm, and}\ s_0 = 0\ \text{mm}\}$.

$$SK = \frac{1}{5} \cdot s\Big|_{R_P = 0.993 \cdot R_{P\infty}} \tag{8.83}$$

In summary, it is necessary and sufficient to measure with the pilot oscillator, as described above, the dependence $R_p$ versus $s$, to plot the measured curve, and to graphically determine three parameters of this curve: $R_{P\infty}$and $s_0$ values as well as the extension of the segment between the distances corresponding to $R_P = 0$ and $R_P = 0.993 \cdot R_{P\infty}$. The constant $SK$ is the fifth part of this segment.

After these evaluation steps, the three values (e.g., in our case $R_{P\infty} = 27.35$ kΩ, $SK = 1.4$ mm, and $s_0 \approx 0$ mm) have to be inserted into (8.82) and the theoretical curve can be represented (see the second curve $RP_{\text{calculated}}$ in Figure 8.18). The comparison of the two plots in Figure 8.18 reflects the accuracy of this modeling.

The benefit of doing this is to get an analytic model of the characteristic $R_P$ versus $s$, which can be imported in SPICE simulation tools to simulate the oscillator operation if the target moves (see Example 8.3). The procedure was applied several times with the author's team and delivered satisfactory results.

Moreover, the natural exponential function has the useful property to be linear in the range limited by the first system constant $SK$. In this subrange, the following dependence is valid:

$$R_P(s) = \frac{0.6321 \cdot R_{P\infty}}{SK} \cdot s \quad \text{for} \quad s \in [0 \ldots SK] \tag{8.84}$$

The model expressed by (8.82) also shows the limits of the system concerning the extension of the sensing range. Theoretically, this is limited to five system constants $SK$ and efforts to increase the range have to be oriented in the enhancement direction of this $SK$ constant by physical and constructive measures.

The two-dimensional graph in Figure 8.18 shows the dependence $RP_{\text{measured}}$ versus distance $s$ for an oscillator, whose maintaining section has a high input impedance and does not parasitically load the resonant LC circuit. This condition is fulfilled by the schematic in Figure 8.17(a).

Unfortunately, this loading is in many cases existent and significantly downgrades the system sensitivity.

The three-dimensional representation in Figure 8.19 shows an array of curves: $RP_{\text{measured}}$ vs. distance $s$, for a distance range between 0 and three system constants $SK$ (see (8.83)) and for an input impedance $R_{\text{INP}}$ of the oscillator electronics varying in the range 10 $R_P$ to 0.1 $R_P$.

Practically, the diagram allows the graphical determination of the system sensitivity at different conditions.

For example, for $R_{INP} = 10\ R_P$, the equivalent parallel resistance $R_P$ drops to 85% from its free-running value (15 % headroom) at a distance, which is equal to 1.6 $SK$. Unfortunately, the sensitivity dramatically decreases so that the distance for 15% headroom reduces to 0.33 $SK$ if $R_{INP} = 0.1\ R_P$.

Figure 8.19 shows the same evaluation using already represented curves for headrooms of 25%, 35%, 50%, and 70%, respectively.

In parallel to the benchmarking of the measured and calculated characteristics $R_P$ versus distance $s$ represented in Figure 8.18, a third database can be taken

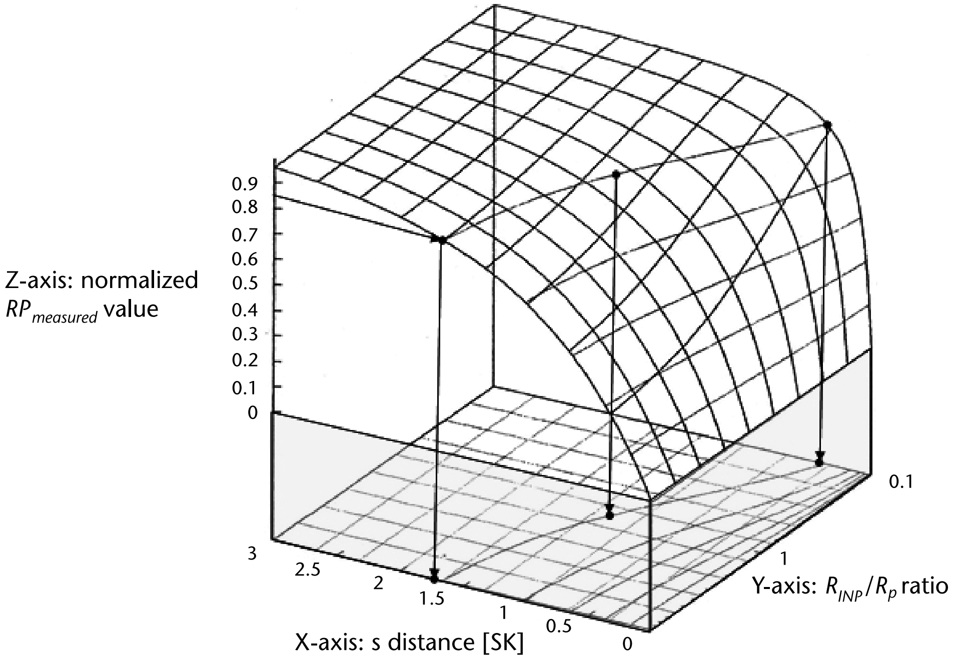


**Figure 8.19** Three-dimensional representation of the $RP_{\text{measured}}$ versus distance $s$, taking into consideration the negative influence of the input resistance $R_{INP}$ of the oscillator electronics.

into consideration, namely the characteristic obtained by means of CAE tools as described in Section 4.3. The field simulation can be repeated at different values of the parameter: distance $s$ (parametric evaluation) and the third characteristic, namely *RPcomputed* is available. Figure 8.20 shows both theoretical characteristics: the computed curve *RPcomputed* beside the calculated curve *RPcalculated* (see (8.82)).

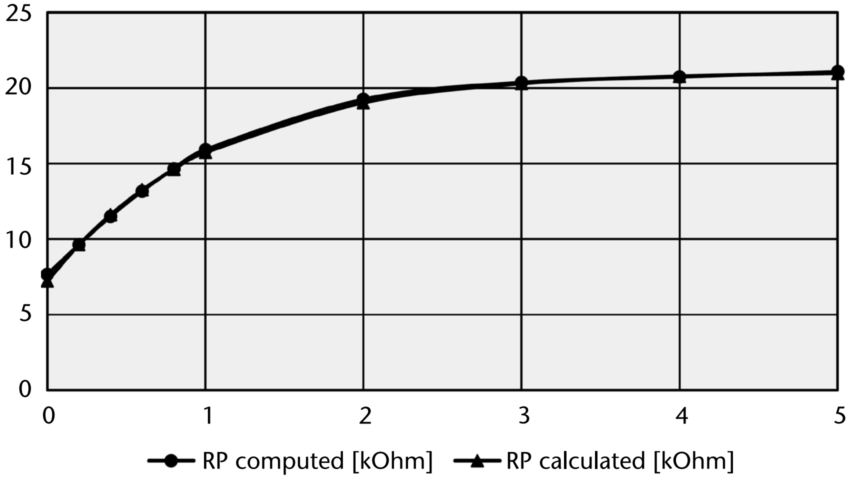


**Figure 8.20** Two versions of the characteristic $R_P$ versus distance $s$, computed with the ANSYS Maxwell electromagnetic field simulator and calculated with (8.82) for {$R_{P\infty}$= 21.07 kΩ, $SK$ = 1.05 mm, and $s_0$ = −0.44 mm}.

Note that another dependence $R_P$ versus $s$ of a different LC dipole was intentionally chosen to give an example with the parameter $s_0$ that is different from zero ($s_0 = -0.44$ mm).

A second more advantageous version of the ring-circuit oscillator is shown in Figure 8.21(a). In contrast to the version in Figure 8.17(a), the parallel resonant LC circuit rests on the real GND, which gives better stability and EMC properties.

The oscillator in Figure 8.21(a )consists of

- A voltage-to-current amplifier (transconductance amplifier) [7] with high input and output impedances, respectively, and a transfer function (transfer admittance):

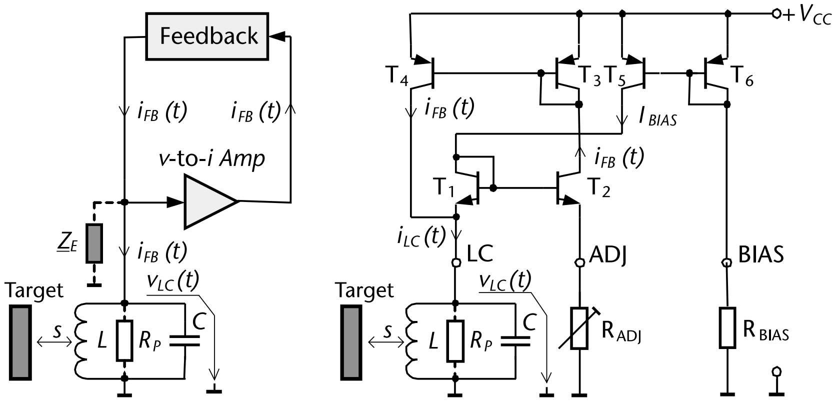
$$\underline{Y}_m \overset{\text{def}}{=} \frac{\underline{I}_{\text{OUT}}}{\underline{V}_{\text{INP}}} = \frac{i_{FB}}{v_{LC}} \tag{8.85}$$

- A current-to-voltage convertor (transimpedance feedback circuit) with a transfer impedance:

$$\underline{Z}_m \overset{\text{def}}{=} \frac{\underline{V}_{\text{OUT}}}{\underline{I}_{\text{INP}}} = \frac{v_{LC}}{i_{FB}} \tag{8.86}$$

First, the oscillator parts are considered ideal, which means the impedance $\underline{Z}_E$ (Figure 8.21(a)), representing the impedance of the electronic stages (input impedance of the amplifier in parallel to the output impedance of the feedback stage), is infinite and can be neglected.

With these definitions, the Barkhausen conditions for gain and phase shift (see (8.57) in Section 8.2.3) are expressed by the following two equations that must be satisfied to sustain oscillations:



(a) (b)

**Figure 8.21** Ring-circuit oscillator with LC dipole on GND: (a) functional block diagram, and (b) example of an implementation in bipolar technology able to be integrated.

$$\left|\underline{Y}_m\left(j2\pi f_0\right)\cdot \underline{Z}_m\left(j2\pi f_0\right)\right| = 1 \tag{8.87}$$

and

$$\angle\underline{Y}_m\left(j2\pi f_0\right) + \angle\underline{Z}_m\left(j2\pi f_0\right) = 0°,\ 360°,\ \text{etc.} \tag{8.88}$$

A very convenient embodiment of this oscillator in Figure 8.21(a) is shown in Figure 8.21(b) [8, 9].The effective amplifier is the transistor $T_2$, which operates as voltage-controlled current source.

The biasing of this main transistor is ensured by the transistor $T_1$, connected as a diode and supplied by the DC current source realized with the current mirror $T_5$ and $T_6$. Supposing an ideal current mirror with an input-output current ratio equal to 1:1, the resistor $R_{\text{BIAS}}$ connected to the oscillator terminal *BIAS* sets the following DC bias current:

$$I_{\text{BIAS}} = \frac{V_{CC} - V_{BE6}}{R_{\text{BIAS}}} \tag{8.89}$$

The level shifter realized with $T_1$ (connected as diode) provides a biasing voltage $V_{BE1}$, which should be as close as possible to the voltage $V_{BE2}$ of $T_2$. A usual value of the DC bias current is $I_{\text{BIAS}} = 100\ \mu\text{A}$ and the typical $V_{BE}$ voltage results between 0.6 to 0.7V [10]. The bias current still flows to the GND through the inductor L. The voltage drop on the DC resistance of the inductor (a few ohms) results below 1 mV and can certainly be ignored.

Under these conditions and assuming an achieved steady state of the oscillations $v_{LC}(t)$, the control voltage of the main transistor is:

$$v_{B2} = v_{LC}(t) + V_{BE1} = V_{lcm}\sin\omega_0 t + V_{BE1} \tag{8.90}$$

For a good matching between $T_1$ and $T_2$ ($V_{BE1} \approx V_{BE2}$), the emitter voltage of $T_2$ closely follows the $v_{LC}(t)$ wave during its positive amplitude and it is zeroed during the negative half-wave ($T_2$ is turned off).

This voltage, available at the oscillator terminal ADJ, is converted into current by the external adjustment resistor $R_{\text{ADJ}}$. The adjusted current is equal to the output current of the voltage-to-current amplifier (transistor $T_2$). This current is actually the feedback current in the oscillator and can be expressed by:

$$i_{FB}(t) = \begin{cases} \dfrac{V_{lcm}\cdot\sin\omega_0 t}{R_{ADJ}} & \text{for } 0 < \omega_0 t < \pi \\ 0 & \text{for } \pi < \omega_0 t < 2\pi \end{cases} \tag{8.91}$$

The Fourier series representation of this positive sine half-wave carries out to a sum of a DC component and more sinusoidal waves with frequencies $f_0$, $2f_0$, $4f_0$, and so forth [5]:

$$
\begin{aligned}
i_{FB}(t) &= \frac{1}{\pi}\cdot\frac{V_{lcm}}{R_{ADJ}} + \frac{V_{lcm}}{2\cdot R_{ADJ}}\sin\left(2\pi f_0 t\right) \\
&\quad - \frac{2\cdot V_{lcm}}{\pi\cdot R_{ADJ}}\left[\frac{\cos\left(4\pi f_0 t\right)}{2^2-1} + \frac{\cos\left(8\pi f_0 t\right)}{4^2-1} + \dots\right]
\end{aligned}
\tag{8.92}
$$

Further on, this current is the input of the feedback stage (second current-to-voltage convertor), which is realized with an AC operating current mirror ($T_3$ and $T_4$). If this second current mirror is similarly an ideal circuit, it has a gain equal to one and does not provoke any phase rotation. Its output current could be considered to be identical to the initial $i_{FB}(t)$ and is finally refed into the parallel resonant LC circuit.

The DC component of the refed current ($V_{lcm}/\pi R_{ADJ}$) has a value in the range of a few tenths of microamps and can also be neglected. Due to the high value of the $Q$ factor and thus of the low value of the bandwidth $B_Q$ (see Table 8.1)—at the bottom of the high selectivity of the parallel resonant LC circuit—the harmonics of the current wave are strongly suppressed (see (8.78)).

The only effective component is the fundamental wave, having the same phase and frequency $f_0$ as the $v_{LC}(t)$ and an amplitude according to the expression (8.92):

$$
I_{fbm} = \frac{V_{lcm}}{2\cdot R_{ADJ}} \tag{8.93}
$$

In summary, this current sets the amplitude of the oscillations and the final oscillations expression results:

$$
v_{LC}(t) = R_P\cdot I_{fbm}\sin\left(2\pi f_0 t\right) = \frac{R_P}{2\cdot R_{ADJ}}\cdot V_{lcm}\cdot\sin\left(2\pi f_0 t\right) \tag{8.94}
$$

The free-running oscillation requirement can be determined making the following simplifying hypotheses: (1) marginal influence of the resonant circuit by target (target absent or far away from to the coil) and (2) negligible impedance of the electronic stages (similar to $\underline{Z}_E$ in Figure 8.21(a)) in parallel to the resonant circuit.

Under these conditions, the current into and the voltage over the resonant circuit are in phase and in this circuit only active power dissipation occurs.

Making an energy balance, this active power, which is dissipated in $R_P$ (current through $R_P$ and voltage over $R_P$ in phase at resonance) during a full time period of the signal $v_{LC}(t)$ must be fed during the positive half-wave of $v_{LC}(t)$. During the negative half-wave the current flowing is switched off at this kind of oscillators.

The dissipated power $P_{DS}$ can be expressed by the integral of the instantaneous power of the wave $v_{LC}(t)$, calculated over an entire period $T_0$:

$$
\begin{aligned}
P_{DS} &= \frac{1}{T_0}\int_0^{T_0} v_{LC}(t)\,i_{LC}(t)\,dt \\
&= \frac{G_P}{T_0}\int_0^{T_0} v_{LC}^2\,(t)\,dt = \frac{G_P V_{lcm}^2}{2\pi}\int_0^{2\pi}\sin^2\omega t\,d\omega t
\end{aligned}
\tag{8.95}
$$

where $G_P = 1/R_P$ is the conductance of the resonant circuit (see (4.38)).

Computing the integral in (8.95) the following expression results in:

$$P_{DS} = \frac{G_P \cdot V_{lcm}^2}{2} \tag{8.96}$$

The delivered power $P_{DV}$ can be computed in a similar way to taking into consideration the effective component of the current $i_{FB}(t)$, namely the fundamental wave with the period $T_0$ (see (8.92)):

$$\begin{aligned} P_{DV} &= \frac{1}{T_0}\int_0^{T_0} \frac{i_{FB}^2(t)}{G_P}\,dt = \frac{1}{T_0 G_P} \cdot \frac{I_{fbm}^2}{4}\int_0^{T_0} \sin^2(t)\,dt \\ &= \frac{I_{fbm}^2}{4G_P \cdot 2\pi}\int_0^{2\pi} \sin^2 \omega t\,d\omega t \end{aligned} \tag{8.97}$$

and after the integral computing:

$$P_{DV} = \frac{I_{fbm}^2}{8 \cdot G_P} \tag{8.98}$$

According to the law of conservation of energy, the delivered energy must be equal to the energy that is converted to heat. Correspondingly, the powers $P_{DS}$ and $P_{DV}$ have to be equal to each other. The equalization of (8.96) and (8.98) leads to the expression of the required amplitude of the feedback current:

$$I_{fbm} = 2G_P \cdot V_{lcm} \tag{8.99}$$

This current originally proceeds from the voltage-controlled current generator, transistor $T_2$, which is controlled by the voltage $v_{LC}(t)$ and is active during the positive half-wave of this voltage, and has the value that is trimmed by the adjusting resistor $R_{ADJ}$:

$$I_{fbm} = \frac{V_{lcm}}{R_{ADJ}} \tag{8.100}$$

The expressions (8.99) and (8.100) finally deliver the adjustment condition of the oscillator, valid for the aforementioned free-running requirements:

$$R_{ADJ} = \frac{1}{2 \cdot G_P} = \frac{1}{2} \cdot R_P \tag{8.101}$$

where the resistor $R_P$ (represented by dashed line in Figure 8.21(a)) symbolizes the parallel loss resistance (see (4.38) for its definition and (5.35) for total $R_S$ expression).

The impedance $\underline{Z}_E$ in Figure 8.21(a), which generally represents the impedance of the electronic stages (input impedance of the amplifier in parallel to the output impedance of the feedback stage), can be estimated for the circuit in Figure 8.21(b) considering its particular equivalent circuit. The impedance $\underline{Z}_E$ is the result of the output resistance $R_{\text{OUT\_34}}$ of the feedback current mirror ($T_3$, $T_4$ ) connected in parallel with the series between the dynamic resistance $R_{D1}$ of the diode $T_1$ and the output resistance $R_{\text{OUT\_65}}$ of the current mirror ($T_6$, $T_5$) of the bias current $I_{\text{BIAS}}$. These magnitudes can sufficiently accurate be approximated by means of the popular formulas used in the design of integrated circuits [10]:

$$R_{\text{OUT\_34}} = \frac{V_A}{\overline{I}_{FB}};\ R_{\text{OUT\_65}} = \frac{V_A}{I_{\text{BIAS}}};\ R_{D1} = \frac{V_T}{I_{\text{BIAS}}} \tag{8.102}$$

where $\overline{I}_{FB}$ is the average value of the current $i_{FB}(t)$, $V_A$ is the Early voltage of the used transistors (typical values: 15 to 100V) and $V_T$ is the thermal voltage, which is well known expressed by:

$$V_T = \frac{k \cdot T}{q} \approx 26 \text{ mV at } 300^\circ\text{K} \tag{8.103}$$

with

$k$ = Boltzmann's constant ($1.380 \cdot 10^{-23}$ joules/kelvin);

$T$ = the absolute temperature in °kelvin;

$q$ = elementary (electron) charge ($1.602 \cdot 10^{-19}$ coulombs).

If the impedance $\underline{Z}_E$ cannot be neglected, it appears in parallel to the $R_P$ equivalent value of the resonant circuit. The lower resulting parallel resistance demands the reduction of the adjustment resistor to get the same amplitude of $v_{LC}$. Concluding, the adjustment condition of this oscillator becomes larger:

$$R_{ADJ} \leq \frac{1}{2} \cdot R_P \tag{8.104}$$

The oscillator current consumption can immediately be expressed by:

$$I_{CC} = 2 \cdot I_{\text{BIAS}} + 2 \cdot \overline{I}_{FB} \tag{8.105}$$

where $\overline{I}_{FB}$ is the average value of the current $i_{FB}(t)$.

*Example 8.2*

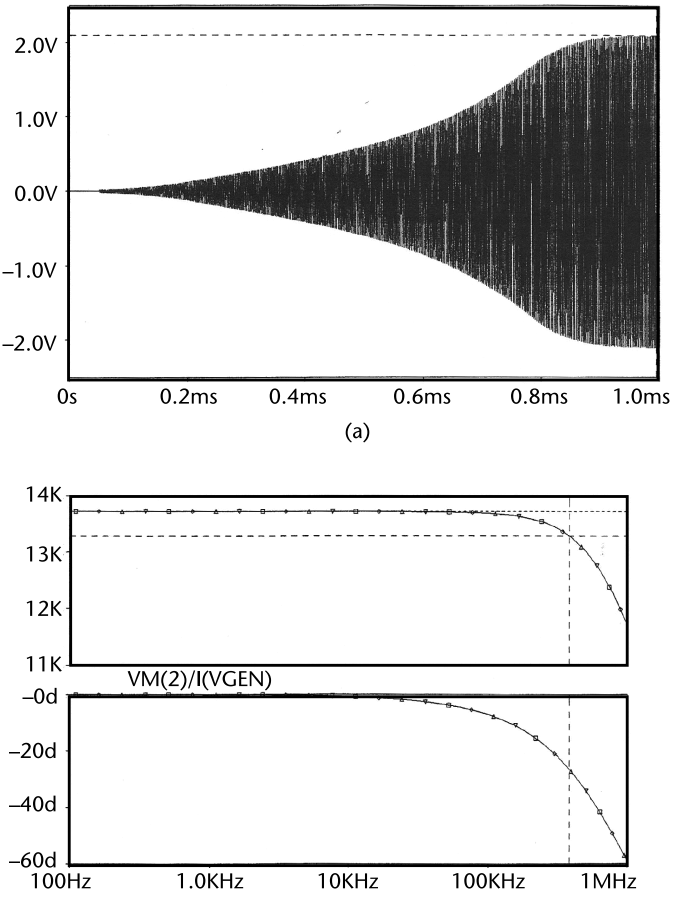
Using the representation (8.92) to express $\overline{I}_{FB}$ and considering the following exemplifying setup: $V_{lcm} = 2\text{V}$, $R_{ADJ} = 10\ \text{k}\Omega$ and $I_{\text{BIAS}} = 100\ \mu\text{A}$, the oscillator current consumption can be estimated by:

$$\begin{aligned} I_{CC} &= 2 \cdot I_{\text{BIAS}} + \frac{2}{\pi} \cdot \frac{V_{lcm}}{R_{ADJ}} \\ &= 200\ \mu\text{A} + \frac{2}{\pi} \cdot \frac{2{,}000\ \text{mV}}{10\ \text{k}\Omega} \approx 327{,}32\ \mu\text{A} \end{aligned} \tag{8.106}$$

Besides the theoretical considerations above and the computation of some features, the perfect evaluation of the oscillator in Figure 8.21(b) is provided by the SPICE simulation tool. To give an example of results, which exclusively can be obtained by circuit simulation, Figure 8.22 shows (a) the startup behavior of the free-running operating mode after the connection of the power supply and (b) the input impedance and its frequency sweep.

This oscillator effectively fulfils the demands listed at the beginning of Section 8.3. It is flexible to allow for adjustment of the working range and it is adequate to be integrated in bipolar technology as well as in CMOS processes. If the circuitry is integrated, RADJ and RBIAS remain external resistors.

The circuit is suitable for parallel resonant LC circuits without tapped coils. It does not need coupling capacitors, which could cause unwanted phase rotations and works from low supply voltages (theoretically higher than 2 $V_{BE}$) up to high voltages (35–60 VDC), has a low current consumption and it is immune enough against environment influences. The accuracy of the input-output ratio of the feedback



(a)

(b)

**Figure 8.22** Ring-circuit oscillator with LC-dipole on GND: (a) the startup behavior of the oscillator output voltage $v_{LC}(t)$, and (b) the magnitude and the phase (units in kΩ and degrees, respectively)of the oscillator input impedance versus frequency (PSpice simulation results).

stage can be improved by replacing the simple current mirror ($T_3$, $T_4$) with a Wilson current mirror [10].

The fundamental principle of a ring-circuit oscillator is widely used in ASICs for inductive sensors.

Well-known companies like Siemens in Germany, Omron Corporation in Japan, and Balluff in Germany [8, 9] have used this schematic core-oscillator for their ASICs for proximity sensors in the last 30 years. Looking at the front-end schematic of the circuit TCA505 (Figure 7.18), a ring-circuit oscillator with a very similar topology is recognizable. The difference consists of the use of a Widlar current mirror [10] with emitter resistor ($T_{124}$, $T_{125}$ and $R_{34}$) instead of the classical one.

The ring-circuit oscillator has been continuously improved and is currently still in use [11, 12].

### 8.3.3 Differential Amplifier Oscillators with LC Dipole

This class of oscillators, also called oscillators with emitter coupling or source coupling, depending on the used implementation technology (bipolar or CMOS), can be considered as the second suitable category of electronic oscillators for inductive sensors besides the ring-circuit oscillators.

Their widespread use is explained by the soft and good controllability of the generated output signal: the amplitude of the AC oscillations is controllable by a DC current and by the smart possibility to define or to compensate the oscillator temperature behavior. The temperature behavior can also be set or compensated via DC signals.

The loading influence on the resonant circuit is very low and the operating frequency range is very large (≤200 MHz). The circuitry is adequate for integration either in bipolar or in CMOS technologies.

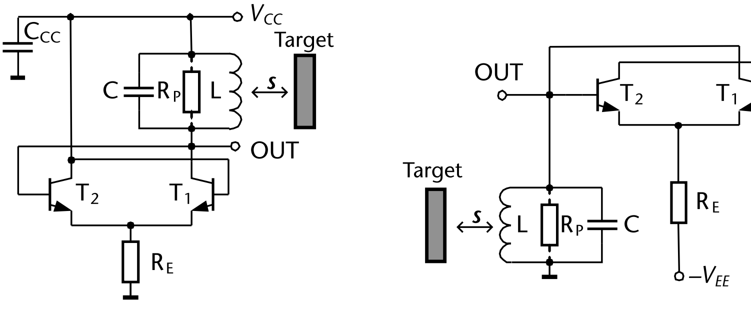
The inductor of the resonant circuit does not have any taps. The LC circuit, which essentially belongs to the inductive sensing element, can be connected either to $V_{CC}$ or—more advantageously—to GND (Figure 8.23).

In both cases, the positive feedback in the differential amplifier with $T_1$ and $T_2$ is provided by the direct connection between two signals in phase, namely the collector of $T_1$ and the base of $T_2$. The open-loop gain is proportional to transconductance $g_m$ of the transistors and can be controlled over a wide range by the emitter current, which means by the emitter resistor $R_E$ [10]. The transistors should be matched to each other; matching pairs are readily obtained by integration techniques.

The schematic in Figure 8.23(a) is integrated in several ICs (e.g., in MC100EL1648, manufacturer ON Semiconductor), which are based on the VCO circuit topology of the classical device MC1648 (manufacturer Motorola Inc.).

The oscillators can be supplied with very low voltages (above 1V) and therefore they are suitable for two-wire inductive proximity sensors as well as for up-to-date schematics with microcontrollers operating at 3.3V.

First improvement steps of the base schematic in Figure 8.23(a) are visible in Figure 8.24. In fact, the emitter resistor $R_E$ is replaced by an emitter biasing current source $2I_{BIAS}$, realized with the current mirror $T_3$, $T_4$ and adjusted by the bias resistor $R_{BIAS}$. In addition, a resistor $R_{FB}$ is looped into the feedback connection: collector of $T_1$ and the base of $T_2$.


Ccc
Vcc
Target
C
RP
L
s
OUT
T2
T1
RE
OUT
T2
T1
Target
s
L
RP
C
RE
−VEE

(a) (b)

**Figure 8.23** Simplified circuit diagrams of the oscillator with emitter coupling: (a) LC dipole rests on $+V_{CC}$, and (b) LC dipole rests on GND.

The startup condition can more easily be determined considering first $R_{FB} = 0$, a good transistor matching, and the steady-state operation mode already installed.

A variation:

$$\Delta v_{BE} = v_{BE1} - v_{BE2} > 0 \tag{8.107}$$

carries out an increase $\Delta i_{C1}$ of the collector current of $T_1$ and, therefore, a change of the collector voltage:

$$\Delta v_{C1} = -R_P \cdot \Delta i_{C1} \tag{8.108}$$

where $R_P$ represents the parallel loss resistance of the resonant circuit (for details, see for example the explanation to (8.101)).

Applying KCL to the emitter circuit to calculate $\Delta i_{C1}$ gives:

$$i_{C1} + i_{C2} = 2I_{\text{BIAS}} \tag{8.109}$$

where the collector currents can be expressed by [10]:

$$\begin{aligned} i_{C1} &= I_{S1} \cdot \left(e^{v_{BE1}/V_T} - 1\right) \\ i_{C2} &= I_{S2} \cdot \left(e^{v_{BE2}/V_T} - 1\right) \end{aligned} \tag{8.110}$$

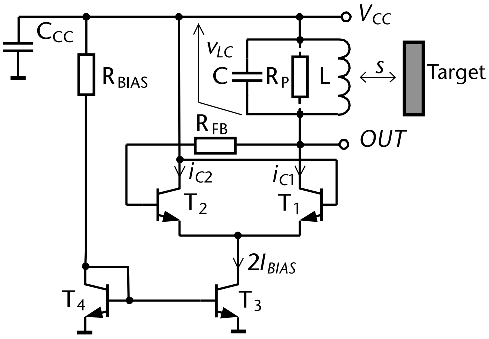

Ccc
Vcc
RBIAS
vLC
C
RP
L
s
Target
RFB
OUT
iC2
iC1
T2
T1
2IBIAS
T4
T3

**Figure 8.24** Relative to the first improvement step of oscillator with emitter coupling.

where $v_{BE1}$ and $v_{BE2}$ are the base-emitter voltages of the transistors $T_1$ and $T_2$, $V_T$ is the thermal voltage [see (8.103)] and $I_{S1}$ and $I_{S2}$ are the corresponding saturation currents. For matched transistors working under identical conditions, the saturation currents are identical.

Dividing $i_{C1}$ by $i_{C2}$ results in an expression for $i_{C2}$:

$$i_{C2} = \frac{1}{e^{\Delta v_{BE}/V_T}} \cdot i_{C1} \tag{8.111}$$

which can be put into (8.109). The new formula for $i_{C1}$ is now:

$$i_{C1} = \frac{2 \cdot I_{\text{BIAS}}}{1 + e^{-\Delta v_{BE}/V_T}} \tag{8.112}$$

For changes $\Delta v_{BE}$ smaller than the thermal voltage, the exponential dependence can be linearized making a Taylor series representation [5]. The collector current becomes:

$$i_{C1} \approx I_{\text{BIAS}} \cdot \left(1 + \frac{\Delta v_{BE}}{2V_T}\right) \tag{8.113}$$

and the increase caused by the variation $\Delta v_{BE}$ results in:

$$\Delta i_{C1} \approx I_{\text{BIAS}} \cdot \frac{\Delta v_{BE}}{2V_T} \tag{8.114}$$

The change $\Delta v_{C1}$ (8.108) is instantly transmitted to the base of $T_2$ as feedback signal $\Delta(v_{BE})^* = -\Delta v_{C1}$. It is in phase with the base voltage of $T_2$ and is expressed by:

$$\Delta\left(v_{BE}\right)^* = -\Delta v_{C1} \approx R_P \cdot \frac{I_{\text{BIAS}}}{2V_T} \cdot \Delta v_{BE} \tag{8.115}$$

In conclusion, the oscillation startup occurs if the gain requirement is also fulfilled; that is, when:

$$R_P \cdot \frac{I_{\text{BIAS}}}{2V_T} = R_P \cdot \frac{g_m}{2} \geq 1 \tag{8.116}$$

Finally, (8.116) gives the setup adjustment formula for this type of oscillator:

$$I_{\text{BIAS}} \geq \frac{2V_T}{R_P} \tag{8.117}$$

The condition is very easy to be implemented by the usual inductive sensors. For $R_P$ values above 5 kΩ, the required emitter current has low values in the range of a few microamps. This current can be adjusted by $R_{\text{BIAS}}$ without a problem:

$$R_{\mathrm{BIAS}} = \frac{V_{CC} - V_{BE}}{2I_{\mathrm{BIAS}}} \tag{8.118}$$

The second change, namely the use of a series feedback resistor ($R_{FB}$ = 100 to 500 Ω) improves the AC behavior as well as the DC current splitting through the transistors. For integrated versions, the resistor can be replaced by a transistor connected as a diode (see $T_7$ in Figure 8.25(a)).

To eliminate the unwanted temperature dependence in the adjustment condition (8.116) caused by $V_T$, the current $I_{\mathrm{BIAS}}$ can be generated instead of $T_3$ by voltage-controlled current source, which is controlled by a bandgap reference [10]. This classical circuit generates a temperature-independent reference voltage but also a voltage proportional to $V_T$. Using the voltage $V_T$ as input for the current source, its output will have the same temperature dependence:

$$I_{\mathrm{BIAS}} = K \cdot V_T \tag{8.119}$$

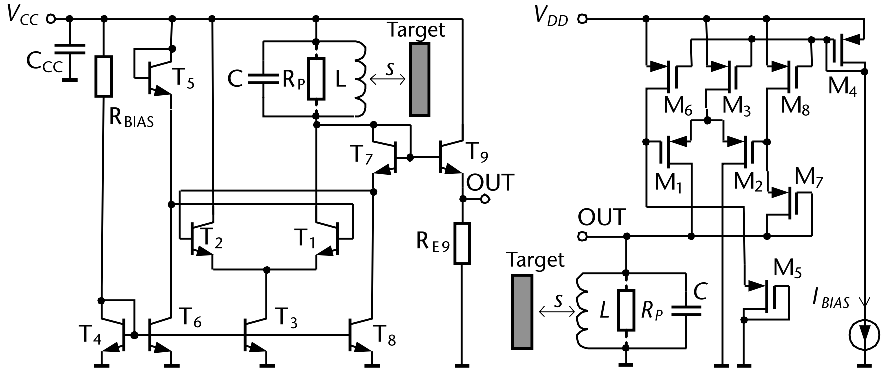
where $K$ is a constant of proportionality.

The substitution of (8.119) into (8.116) leads to an adjustment formula without any temperature-dependent quantities and thus valid independent of temperature:

$$R_P \cdot \frac{K}{2} \geq 1 \tag{8.120}$$

The circuit in Figure 8.24 can furthermore be optimized for an integration under optimal key premises, such as chip area reduction using electronic active parts instead of area-consuming elements (e.g., silicon resistors) and implementation of high-performance elements (Figure 8.25).

The schematic in Figure 8.25(a) has the following add-ons in comparison to the previous version in Figure 8.24:



(a) (b)

**Figure 8.25** High-performance oscillators with emitter coupling for inductive sensors suitable for ASIC implementations: (a) bipolar version with LC dipole connected to $+V_{CC}$ and (b) CMOS implementation with LC dipole connected to GND.

- The circuit exclusively contains npn-transistors, which are superior to integrated pnp-types.
- The feedback resistor $R_{FB}$ is replaced by the diode realized with $T_7$.
- To get a larger excursion of the output signal $v_{LC}(t)$, the base of $T_1$ is supplied one $V_{BE}$ below $V_{CC}$ by means of $T_5$ also connected as diode.
- The biasing current sources for the original differential amplifier ($T_1$ and $T_2$) and also for the level shifter ($T_7$ and $T_5$) are combined into a multioutput current mirror (($T_4$, $T_3$, $T_6$, and $T_8$) that is referenced by the common adjusting resistor $R_{\mathrm{BIAS}}$. The currents are matched by means of suitable emitter areas of the corresponding transistors.
- The emitter follower with $T_9$ reduces the oscillator output impedance, making it closer to an ideal voltage source.

The schematic in Figure 8.25(b) illustrates a similar topology, which is obtained by an overturning of the schematic sketch in Figure 8.25(a) with respect to a virtual horizontal line and by replacement of the npn transistors by pmos transistors with enhancement channel [10].

In fact, the oscillator circuit with emitter coupling is a significant oscillator for inductive sensors. It was the core oscillator version in the TDE ASIC family of the former SGS-Thomson Company and it is still the main oscillator in the TDA family of the resulting company STMicroelectronics. Excellent features, primarily controllability, make this the preferred schematic for many experts in this field [13].

### 8.3.4 Bridge-Network Oscillators

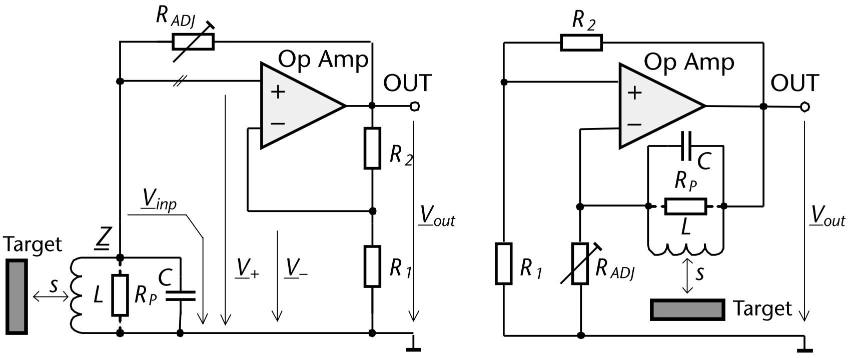
Oscillators with bridge-like networks as frequency-determining section and op-amps as maintaining section (for these definitions see Section 8.2.3) can be considered as forward-looking solutions for inductive sensors. The benefits of this topology are the ideal excitation conditions of the feedback network with minimal parasitic loading through the amplifier and large variation range of the output signal amplitude, which increases the EMC sensor immunity. In addition, the oscillator and the following signal processing (linearization, compensation, etc.) can be consistently implemented using op amp throughout.

Despite these advantages, the need of a symmetrical power supply that is not compliant with the standards for industrial applications, which demands a single supply voltage (see Section 2.1.1), has long impeded the widespread of this solution. Today, the extensive availability of rail-to-rail op-amps brings new life to bridge-network oscillators.

Two versions relating to the connections of the resonant LC circuit are known (Figure 8.26).

The functional block diagram in Figure 8.26(a) [7] shows a bridge-like circuitry around the op-amp, consisting of

- The impedance $\underline{Z}$ of the resonant circuit;
- The adjustment resistor $R_{ADJ}$;
- The voltage divider ($R_1$ and $R_2$) that sets the closed-loop gain $A_F$ of the op-amp (see (8.45)).



(a) (b)

**Figure 8.26** Functional block diagram of the bridge-network oscillator: (a) resonant LC circuit in the positive feedback loop, and (b) resonant LC circuit in the negative feedback loop.

Additionally, a coupling capacitor not illustrated in Figure 8.26(a) and having a capacitive reactance negligible in comparison with $R_{ADJ}$ could be connected in series with $R_{ADJ}$ to prevent a DC-positive feedback.

Working in the complex *s*-domain and considering an ideal op-amp (see its definition in Sections 8.3.1 and 8.3.2), the complex magnitude of the output voltage can be expressed by:

$$\underline{V}_{\text{out}} \triangleq \left(\underline{V}_{+} - \underline{V}_{-}\right)A_F = \frac{R_{ADJ} + \underline{Z}}{\underline{Z}}\underline{V}_{\text{inp}} \tag{8.121}$$

where the input voltage at the inverting terminal of the op-amp is:

$$\underline{V}_{-} = \frac{R_1}{R_1 + R_2}\underline{V}_{\text{out}} \tag{8.122}$$

The substitution of (8.122) into (8.121) and some computation steps lead to an expression of the loop gain $T$ (see (8.46):

$$T(s) \overset{\text{def}}{=} \frac{V_{\text{inp}}}{V_{+}} = -\beta(s)A(s) = \frac{A_F \cdot \underline{Z}}{\left(R_{ADJ} + \underline{Z}\right) \cdot \left(1 + A_F \cdot \dfrac{R_1}{R_1 + R_2}\right)} \tag{8.123}$$

For high values of the gain with feedback $A_F$, the expression (8.123) can be simplified:

$$T(s) \approx \frac{R_1 + R_2}{R_1 \cdot \left(1 + R_{ADJ} \cdot \underline{Y}\right)} \tag{8.124}$$

where $\underline{Y}$ is the complex admittance of the resonant circuit (see (4.13)). As before, due to the parallel topology of this LC circuit, the use of the complex admittance is preferred to the complex impedance. Its expression is:

$$\underline{Y} \overset{\text{def}}{=} \frac{1}{\underline{Z}} = \frac{1}{R_P} + \frac{1}{sL} + sC \tag{8.125}$$

with $R_P$ the equivalent parallel loss resistance of the resonant LC-circuit working at its resonance frequency (for details, see for example the explanation to (8.101)).

The feedback factor, which represents the denominator of the system transfer function (frequency response), is expressed by:

$$D(s) = 1 - T(s) \approx 1 - \frac{R_1 + R_2}{R_1 \cdot \left(1 + R_{ADJ} \cdot \underline{Y}\right)} \tag{8.126}$$

Finally, the poles of the system are the roots of the equation:

$$D(s) = 1 - \frac{R_1 + R_2}{R_1 \cdot \left(1 + R_{ADJ} \cdot Y\right)} = 0 \tag{8.127}$$

The substitution of (8.125) into (8.127) yields the final equation for the poles calculation:

$$s^2 \cdot CR_{ADJ} + s \cdot \Delta + \frac{R_{ADJ}}{L} = 0 \tag{8.128}$$

where the operand Δ is:

$$\Delta = \frac{R_{ADJ}}{R_P} - \frac{R_2}{R_1} \tag{8.129}$$

The poles of the system are expressed by the well know solution:

$$s_{1,2} = \frac{-\Delta \pm \sqrt{\Delta^2 - 4 \cdot R_{ADJ}^2 \cdot \frac{C}{L}}}{2 \cdot C \cdot R_{ADJ}} \tag{8.130}$$

As Section 8.2.3 explained, the system has started oscillation if the real parts of the poles are:

$$\Re\left\{s_{1,2}\right\} = -\Delta > 0 \tag{8.131}$$

and the steady-state oscillations are obtained if:

$$\Re\left\{s_{1,2}\right\} = -\Delta = 0 \tag{8.132}$$

The condition (8.132) leads to the adjustment requirement:

$$R_{ADJ} = R_P \cdot \frac{R_2}{R_1} \tag{8.133}$$

where $R_P$ is the equivalent parallel loss resistance of the resonant circuit working at its resonance frequency.

The condition (8.133) yields the name bridge-network oscillator.

As expected, the imaginary part resulting from (8.125) for $\Delta = 0$ leads to the Thomson's formula:

$$\omega = \Im\left\{s_{1,2}\right\}\Big|_{\Delta=0} = \frac{1}{\sqrt{LC}} \tag{8.134}$$

Figure 8.26(b) shows the second version of the bridge-network oscillator, in which the resonant circuit is connected between the output of the op-amp and its inverting input, while the voltage divider $R_1$ and $R_2$ provides the positive feedback path [14]. The operation mode is very similar to the first version in Figure 8.26(a) and does not need additional information.

### 8.3.5 Oscillators with Pulsing DC Current Excitation

According to professional and patent publications [15], this type of oscillator is predominantly used in inductive sensors working as displacement sensors. The explanation results from the specific operating mode: their output signal is proportional to the equivalent parallel resistance $R_P$ (see (4.38) for its definition and (5.35) for total $R_S$ expression). In this way, the signal very closely follows the dependence of $R_P$ on the distance $s$ between the target and the coil in the inductive sensing element (see the plot called RPmeasured in Figure 8.18).

As expressed by (8.84), this dependence has a good linearity in the range from $s = 0\%$ up to 63% of its extension, namely:

$$R_P(s) = \frac{0.6321 \cdot R_{P\infty}}{SK} \cdot s \text{ for } s \in [0\ldots SK] \tag{8.135}$$

For the quantities in (8.135), see (8.84).

The basic circuit of this oscillator in Figure 8.27(a) shows a pulsing DC current source, which supplies the resonant parallel LC circuit and is—at the same time—controlled by this circuit. The parallel resonant LC circuit, which belongs to the inductive sensing element, is periodically supplied from the constant DC current source $I$ through a high-speed analog switch SW. The switch on/switch off command of the SW is provided by the amplifier A with a high-impedance input. The command is synchronous with the voltage $v_{LC}(t)$ over the LC circuit.

The formula for the signal amplitude, given by the patent holder Electro Corporation, Sarasota, Florida, in U.S. patent 4,001,718:

$$V_{lcm} = K \cdot I \cdot Q_L \cdot X_{L0} \tag{8.136}$$

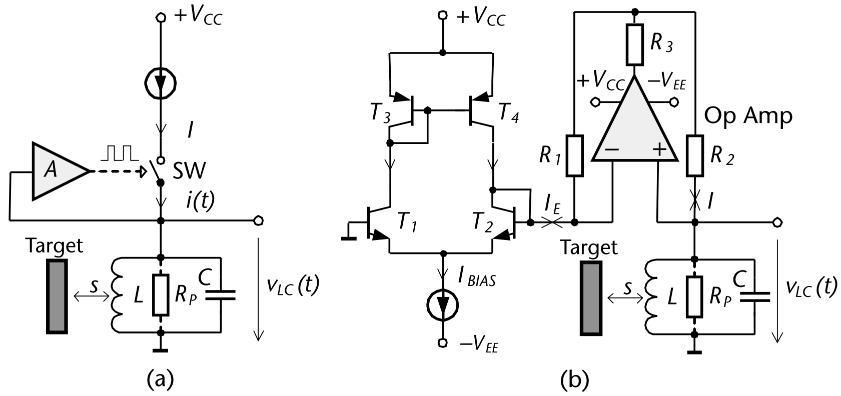


**Figure 8.27** Oscillator with pulsing DC current excitation: (a) basic circuit with analog switch (SW), and (b) version with bidirectional DC current source.

where $K$ is a multiplication constant, $Q_L$ is the $q$ factor of the coil (see Table 4.1), and $X_{L0}$ is the coil reactance at the resonance frequency $\omega_0$, can immediately be converted to:

$$V_{lcm} = K \cdot I \cdot \frac{\omega_0 L}{R_S} \cdot \omega_0 L = K \cdot I \cdot R_P \tag{8.137}$$

This expression emphasizes the proportionality mentioned above.

The computation of the constant $K$ can be made under ideal conditions; that is, the switch operates without time delay and has a zero switched-on resistance and an infinite switched-off resistance.

The current wave $i(t)$ after the SW is a rectangular signal with the levels 0 and $I$ and with the on/off ratio equal to 1:1. Thus, it can be expressed by the following Fourier series representation [5]:

$$i(t) = I \cdot \left\{ \frac{1}{2} + \frac{2}{\pi} \cdot \left[ \sin\left(2\pi f_0 t\right) + \frac{\sin 3 \cdot \left(2\pi f_0 t\right)}{3} + \frac{\sin 5 \cdot \left(2\pi f_0 t\right)}{5} \right] \right\} \tag{8.138}$$

Similar to (8.78), the DC component and the harmonics of the current wave are strongly suppressed. The effective component is only the fundamental wave, having the frequency $f_0$ and the amplitude:

$$I_m = \frac{2}{\pi} \cdot I \tag{8.139}$$

Hence, the multiplication constant is $K = 2/\pi$ and the impedance-voltage conversion provided by the system is finally expressed by:

$$V_{lcm} = \frac{2}{\pi} \cdot I \cdot R_P \tag{8.140}$$

In order to prevent the use of analog switches and to allow the circuit integration, the schematic can be realized with a bidirectional DC current source (Figure 8.27(b)) [14]. This source can be implemented with a differential amplifier ($T_1$ and $T_2$) having a current mirror ($T_3$ and $T_4$) as an active load. The biasing is provided by an emitter current source, which delivers the constant DC current $I_{\text{BIAS}}$.

Considering an ideal op-amp, the voltage at its inverting input is equal to the voltage at its noninverting input, which is $v_{LC}(t)$.

During the positive half-wave of $v_{LC}(t)$, $T_2$ goes ON and turns off the transistor $T_1$, $T_3$, and $T_4$. A current $I_E$ flows now through the path: the output of the op-amp, resistor $R_1$, BE junction of $T_2$, and emitter current source. Its expression is:

$$I_E = I_{\text{BIAS}} \tag{8.141}$$

During the negative half-wave of $v_{LC}(t)$, $T_2$ goes OFF and $T_1$, $T_3$, and $T_4$ turn on. The bias current $I_{\text{BIAS}}$ is redirected and flows now through the path: $T_1$, $T_3$ and $T_4$, and $R_1$ to the output of op-amp. Consequently, the current $I_E$ flows in the opposite direction and can be expressed by:

$$I_E = -I_{\text{BIAS}} \tag{8.142}$$

Due to the small commutation range of the differential amplifier, the current $I_E$ has a regular rectangular shape. This current wave flowing through the resistor $R_1$ generates a voltage drop, which is equal to the voltage drop on the resistor $R_2$, if $R_1 = R_2$. For ideal op-amp conditions, the current through $R_2$ is the same.

In summary, the supply current of the resonant LC circuit is also rectangular, has the frequency of $v_{LC}(t)$, and is in phase with this voltage. Its value is:

$$I = \begin{cases} I_E = I_{\text{BIAS}} & \text{for } v_{LC}(t) > 0 \\ I_E = -I_{\text{BIAS}} & \text{for } v_{LC}(t) < 0 \end{cases} \tag{8.143}$$

The basic circuit can be extended using an alternating DC current supply of the resonant LC circuit [16]. Hence, the LC circuit is connected to the generator both sides via two switches.

### 8.3.6 Negative-Resistance Oscillators

An alternative form of a maintaining section (see the definition in Section 8.2.3) is a source of negative resistance. This paradoxical name refers to an electronic device or electronics circuit for which a decrease in applied voltage brings about an increase in current through it. There are a number of semiconductor devices or circuits with such voltage-current characteristic having a segment with negative dynamic resistance ($\Delta u/\Delta i < 0$) [2, 6, 17].

#### 8.3.6.1 Negative Resistance: Fundamentals

In the systems with friction (see Section 8.2.2), the frictional coefficient $b$ (see (8.26)) or the corresponding damping coefficient $h = R/2L$ in (8.29) are always positive quantities since friction always hinders motion and the electrical resistance is a positive quantity, respectively [3]. A positive coefficient of friction and a positive resistance mean that in order to overcome frictional forces or resistances in electric circuits, energy has to be spent (see Section 8.2.3).

Starting with (8.26), multiplying all terms by *dx*/*dt*, taking the integral over a certain interval of time from 0 to $\tau$ and carrying out the middle term, the resulting expression:

$$\left|\frac{m}{2}\left(\frac{dx}{dt}\right)^2\right|_0^\tau + \left|\frac{kx^2}{2}\right|_0^\tau = -\int_0^\tau b\left(\frac{dx}{dt}\right)^2 dt \tag{8.144}$$

shows on the left-hand side the variation of kinetic and potential energy of the system during the time from 0 to $\tau$. The sum determines the variation of the total energy of the system over this interval of time.

If $b$ is positive, the integral on the right-hand side is positive and the variation of the total energy results in a negative; that is, the system energy decreases. This is caused by losses of energy owing to friction or to heating dissipation in the electrical circuit.

If $b$ or $h$ were negative, according to (8.144), the system energy would increase and the friction, in this case, would be a source of energy. If the system possesses its own reservoir of energy, then it can be admitted that $b < 0$ or $h < 0$ and that the energy of the system increases at the expense of friction or resistance, even though there is no longer friction or resistance in the usual sense.

Since they are characterized by the same term in the differential equation, namely the term containing *dx*/*dt*, it is usual to also employ the terms friction or resistance for $b < 0$ or $h < 0$, and to speak of negative friction or negative resistance. Reference [3] intuitively shows how mechanical systems with friction can be extended to become negative friction systems and to oscillate. This mechanism is applicable also for electronic oscillators and some implementations are described in the next sections.

In conclusion, the way to obtain a system with negative resistance is to achieve a differential equation of the system having the general expression:

$$\ddot{x} + 2h\dot{x} + \omega_0^2 x = 0 \tag{8.29}$$

and to manipulate the parameter $h$ by means of additional components or circuits in order to make its value to negative.

In order to determine the portrait on the phase plane of the oscillator with negative resistance, it is sufficient to consider the portrait established earlier for the oscillator with losses (see Figure 8.8) and to see how this changes when the sign of $h$ alternates.

Similar to systems with losses, which could have small or large friction, the systems with negative resistance have different portraits depending on the amount of the negative $h$ parameter.

Electronic oscillators with negative resistance are generally characterized by negative friction that is not too large; that is, they fulfill the requirement:

$$h^2 < \omega_0^2 \tag{8.145}$$

The phase path of the oscillators with negative resistance can be equally obtained once more representing (8.35) and (8.36) in the phase plane under the changed conditions (Figure 8.28).

The resulting portrait looks similar. However, now $h < 0$ and the direction of the spirals is changed into the opposite one with respect to that obtained for systems with $h > 0$. Consequently, moving along an integral curve, the representation point will move away from the state of equilibrium (the singular point $x = 0$, $y = 0$). This singular point is also, in this case, the asymptotic point of the spirals family within each other (i.e., it is a singular point of focus type). Due to the fact that $h < 0$, the only position of equilibrium is instable; the singular point is an unstable focus.

The velocity of the motion of the representation point on the phase plane reduces as in previous case to zero at the origin of the coordinates only and increases together with the distance of the representation point from the origin.

The case of systems with large negative friction has a lower importance for electronic oscillators. Reference [3] presents the progressive changeover of the spiral integral curves into a family of integral curves of the parabolic type, all curves passing through the unique singular point situated at the origin of the coordinates. This is a singular point of the node type.

The motion of the system in cases $h < 0$ and (8.145) is also an oscillatory process, just for small positive $h$, but the process is no longer a damped but a reinforcing one. The maximum deviations (amplitude) of the system theoretically increase with the time according to the expression:

$$x = Ke^{-ht} \cdot \cos(\omega t + \alpha) \text{ where } h < 0 \tag{8.146}$$

The law of the amplitude increase is a geometrical progression with the common ratio $e^{-hT}$. The quantity $-hT$ is called a logarithmic increment of the oscillations and since $h < 0$ then $e^{-hT} > 1$.

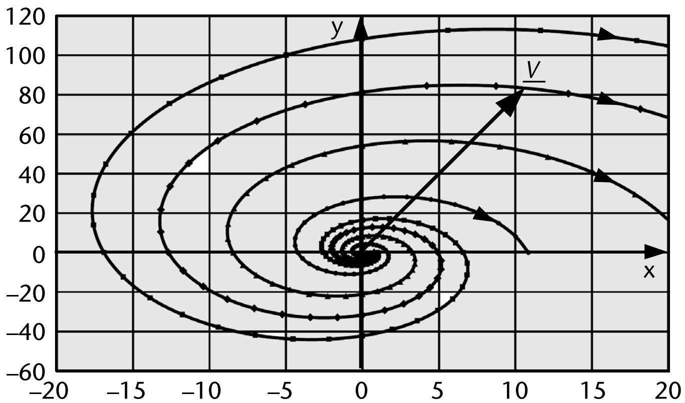


**Figure 8.28** Phase paths of the LC oscillator with negative resistance.

#### 8.3.6.2 Basic Circuit Diagrams of the Negative-Resistance Oscillators

The equivalent circuit diagram in Figure 8.29(a) illustrates a parallel resonant LC circuit with losses (represented by the equivalent parallel resistance $R_P$) and a parallel negative resistance $R_N$, which can be provided by a physical device (e.g., tunnel diode) or obtained by feedback loop(s) in an electronic circuitry (see the following sections).

In Figure 8.29(b) the resistances are replaced by the corresponding conductances. In addition, the network contains the conductance of the maintaining section $G_{MS}$.

Applying KCL to this equivalent network, expressing the partial currents by means of KVL and rearranging the terms yields:

$$\dot{v}_{LC}(t) + \frac{1}{LC}\int v_{LC}(t)\,dt + \frac{G_P + G_N}{C}\cdot v_{LC}(t) = 0 \tag{8.147}$$

where $G_P$ and $G_N$ correspond to $R_P$ and $R_N$, respectively. $G_{MS}$ is neglected.

Carrying out a derivative with respect to the time of this equation, a quadratic differential equation results in:

$$\ddot{v}_{LC}(t) + \frac{G_P + G_N}{C}\cdot \dot{v}_{LC}(t) + \frac{1}{LC}\cdot v_{LC}(t) = 0 \tag{8.148}$$

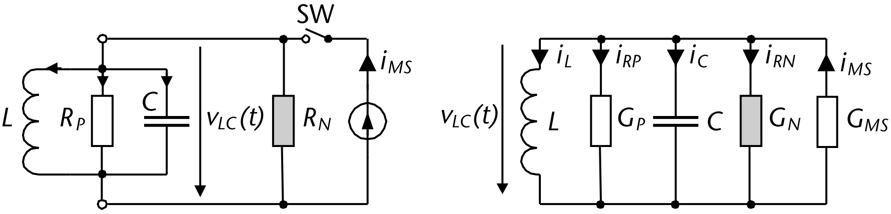
The solution of this equation can be generally expressed by:

$$v_{LC}(t) = V_{lcm}e^{-\frac{t}{2C}\cdot\left(\frac{R_P+R_N}{R_P\cdot R_N}\right)}\cdot\sin(\omega t + \alpha) \tag{8.149}$$

where the amplitude $V_{lcm}$ is dependent on the initial current conditions and the angular frequency $\omega$ is expressed by:

$$\omega = \omega_0\cdot\sqrt{1 - \frac{L}{4C\cdot\left(\frac{R_P\cdot R_N}{R_P + R_N}\right)^2}} = \frac{1}{\sqrt{LC}}\cdot\sqrt{1 - \frac{L}{4C\cdot\left(\frac{R_P\cdot R_N}{R_P + R_N}\right)^2}} \tag{8.150}$$

The following cases are possible depending on the relationship between $R_P$ and $R_N$:



(a) (b)

**Figure 8.29** (a) Basic circuit diagram of the oscillator with negative resistance, and (b) equivalent network.

- $|R_N| > |R_P|$ → The equivalent value of $R_N$ in parallel to $R_P$ is positive, the time constant of the natural exponential function becomes negative, and the oscillations are damped.
- $|R_N| < |R_P|$ → The time constant of the exponential function changes the sign and becomes positive. The oscillations are no longer damped but theoretically continuously increase.
- $|R_N| = |R_P|$ → The negative resistance compensates the losses, the time constant of the natural exponential function is equal to zero, and the oscillator is in a steady-state mode.

This case is important for oscillator application and can be evaluated by solving (8.148) for the equality $|G_N| = |G_P|$. The oscillator differential equation is now given by:

$$\ddot{v}_{LC}(t) + \frac{1}{LC} \cdot v_{LC}(t) = \ddot{v}_{LC}(t) + \omega_0^2 \cdot v_{LC}(t) = 0 \tag{8.151}$$

The solution of this particular differential equation has the general form:

$$v_{LC}(t) = e^{at} \cdot (A\cos bt + B\sin bt) \tag{8.152}$$

where $A$ and $B$ are integration constants. The quantities $a$ and $b$ are the real and imaginary parts of the roots of the algebraic characteristic equation of (8.151):

$$x^2 + \omega_0^2 = 0 \tag{8.153}$$

The roots:

$$x_{1,2} = \pm\sqrt{-\omega_0^2} = \pm j\omega_0 \tag{8.154}$$

have real parts equal to zero and imaginary parts equal to $\omega_0$, so that the quantities $a$ and $b$ mentioned above are:

$$a = \Re\{x_{1,2}\} = 0 \text{ and } b = \Im\{x_{1,2}\} = \omega_0 \tag{8.155}$$

The substitution of these values into the expressions of the voltage $v_{LC}$ (formula (8.152)) and its derivative leads to:

$$v_{LC}(t) = A\cos\omega_0 t + B\sin\omega_0 t \text{ and} \tag{8.156}$$

$$\dot{v}_{LC}(t) = -\omega_0 A\sin\omega_0 t + \omega_0 B\cos\omega_0 t \tag{8.157}$$

The equation system (8.156) and (8.157) gives the constants of integration $A$ and $B$ with respect to the initial conditions.

At initial time, the capacitor C impedes a step variation of the signal $v_{LC}(t)$. This given fact allows determination of the constant $A$:

$$v_{LC}(t)\Big|_{t=0} = 0 \Rightarrow A = 0 \tag{8.158}$$

On the other hand, the initial current cannot start through the inductor; its path leads through the capacitor C. This results in a second initial condition, which permits to express the second constant $B$:

$$\dot{v}_{LC}(t)\Big|_{t=0} = \frac{I_0}{C} \Rightarrow B = \frac{I_0}{\omega_0 C} \tag{8.159}$$

where $I_0$ is the initial current impulse delivered by the electronic maintaining section of the oscillator.

The substitution of (8.158) and (8.159) into (8.156) leads to the final expression of the signal of the oscillator with negative resistance:

$$v_{LC}(t) = \frac{I_0}{\omega_0 C}\sin\omega_0 t = I_0 \cdot \sqrt{\frac{L}{C}} \cdot \sin\frac{t}{\sqrt{LC}} \tag{8.160}$$

As previously mentioned, $R_N$ can be provided by a physical device (see Sections 8.3.6.3 and 8.3.6.4) or obtained by feedback loop(s) in an electronic circuitry (see Section 8.3.6.5).

The use of electronic devices with negative dynamic resistance in oscillators for inductive sensors is not very spread since their v-to-i characteristics are unstable and present a large sample deviation. In addition, the controllability of such oscillators is limited.

Synthesizing methods of negative resistance by means of feedback loops is the popular option [18], [19]. The negative resistance can supplementarily be represented in the oscillator schematic and the problem solving can be performed using the classical procedures for the positive feedback (see the previous sections).

To shape the basic circuit arrangement providing negative resistance, let us recap the block diagram of an amplifier with selective positive feedback (Section 8.2.3). The lossy parallel resonant LC-circuit in Figure 8.30(a) is the frequency-determining section and the amplifier A is the maintaining section of the oscillator.

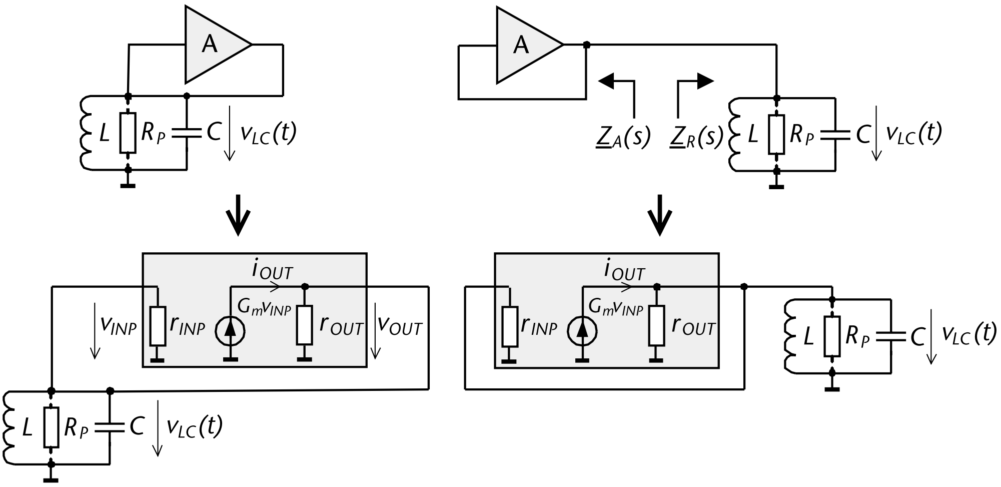
The equivalent circuit of v-to-i amplifier A is represented by the input and output resistances $r_{INP}$ and $r_{\mathrm{OUT}}$ as well as by transconductance $G_m$ defined as [7]:

$$G_m \overset{\text{def}}{=} \frac{i_{\mathrm{OUT}}}{v_{\mathrm{INP}}} \tag{8.161}$$

If a frequency exists at which the magnitude of the loop gain $T$ is greater than one and the total phase shift through the basic amplifier and the feedback network is 360° (or $2\pi n$ radians, where $n$ is a natural number $n = 1, 2, \ldots$), oscillations with this frequency occur.

From a system-theory perspective, that means the need for a complex conjugate pole-pair on the right-hand side of the complex $s$-plane (Figure 8.11).

In the time domain, the result is oscillations with increasing amplitude. The initial stimulus can be the thermal noise, which is already present in electronic



(a) (b)

**Figure 8.30** (a) Amplifier with selective positive feedback (block schematic and equivalent circuit), and (b) alternative approach.

devices. The amplitude increases until the high limit of the dynamic range of the amplifier is reached. Consequently, the amplitude will be maintained limited and the loop gain $T$ decreases returning to one. The poles move toward the imaginary axis and definitively remain on this ($s = \pm j\omega$). The final result is undamped oscillations with constant amplitude.

In addition, this amplitude can be adjusted by an automatic amplitude control: the amplitude is measured, compared with a reference magnitude, and reduced to a value below the high limit of the dynamic range by means of a reduction of the loop gain.

To specify the requirements that negative-resistance oscillators have to fulfill, the circuit in Figure 8.30(a) can be outlined by means of two complex impedances (Figure 8.30(b)) [7]; namely the impedance of the resonant circuit $\underline{Z}_R(s)$ and impedance of the amplifier $\underline{Z}_A(s)$:

$$\underline{Z}_A(s) = \frac{1}{\dfrac{1}{r_{\text{INP}}} + \dfrac{1}{r_{\text{OUT}}} + G_m(s)} \tag{8.162}$$

For a negative transconductance $G_m$ having sufficiently high magnitude, the real part of $\underline{Z}_A(s)$ results in a negative and acts as negative resistance, which compensates the equivalent loss resistance $R_P$ of the resonant circuit.

If a frequency exists at which the impedance $\underline{Z}$(s), defined by:

$$\underline{Z}(s) = \underline{Z}_A(s) + \underline{Z}_R(s) \tag{8.163}$$

fulfills the following demands:

$$\Re\{\underline{Z}(s)\} = \Re\{\ \underline{Z}_A(s) + \underline{Z}_R(s)\} < 0 \text{ and} \tag{8.164}$$

$$\Im\left\{\underline{Z}(s)\right\} = \Im\left\{\underline{Z}_A(s) + \underline{Z}_R(s)\right\} = 0 \tag{8.165}$$

oscillations with this frequency do appear.

During the startup time period, the oscillation's amplitude tends to increase. This causes a movement of $\Re\{\underline{Z}(s)\}$ toward zero and—at the end of the startup time period—an undamped oscillation.

#### 8.3.6.3 Oscillators with Electronic Devices with Negative Resistance

These oscillators need only two connections to the frequency-determining section and thus are very suitable for series or parallel resonant LC circuits. The devices with negative resistance have voltage-current characteristics with a letter S shape or a letter N shape and with following orientations:

- Vertical orientation: A particular value of applied voltage could produce two or three current values while a particular value of applied current could produce only one voltage value. The S-shaped characteristic is called voltage-controlled and is more adequate for series tuned circuits.
- Horizontal orientation: A particular value of applied current could produce two or three voltage values while a particular value of applied voltage could produce only one current value. This much more familiar N-shaped characteristic is called current-controlled and is more adequate for parallel tuned circuits.

In both, the negative-resistance region is confined to a limited range of voltage and current. The negative resistance $r_{\text{diff}}$ is a differential quantity; that is, it is the ration of a small change in voltage to the resulting change in current (small signal mode).

To use such devices in an oscillator, it is sufficient to offset the positive loss resistance of the resonant circuit connected to it. Therefore, the device with negative resistance should have a negative resistance, which is numerically greater than the positive loss resistance of the series resonant circuit or is numerically less than the positive loss resistance of the parallel resonant circuit (see (8.149)). The oscillation amplitude is limited to the value of the average slope of the characteristic part, used in the oscillation process, which corresponds to the equivalent resistance of the resonant circuit [6].

The oscillators with electronic devices with negative resistance have the following features:

- The output signal is very device-specific and generally has small values;
- The oscillator controllability is also restricted, which is a disadvantage for the use in inductive sensing elements;
- The simple oscillator arrangement and the capability to work at medium to high frequencies (up to gigahertz range) are advantageous.

There are more such active dipoles with this characteristic, such as tunnel diode, impatt diode, and trappat diode [2, 6]. The tunnel diode [20] and an emulation circuit called lambda diode [21] have a greater relevance for this book.

The tunnel diode is an HF semiconductor device [2] and was first reported by Esaki (Japan) in 1958. For this reason, it is also called the Esaki diode. It is a heavily doped p-n germanium or silicon diode with a very thin junction region. For this reason, the electrical field strength in the junction region reaches high values ($10^6$ V/cm) and a tunnel movement of the electrons can occur under certain conditions. The v-i characteristic has a region with negative dynamic resistance, which is caused by penetration of the potential barrier at the junction by electrons with insufficient energy to surmount the barrier. The electrons movement through the tunnel occurs with the light velocity. The effect occurs at low forward bias, typically between 0.1 and 0.3V. The breakdown occurs at a very low value of the reverse bias (i.e., there is no region of high reverse resistance).

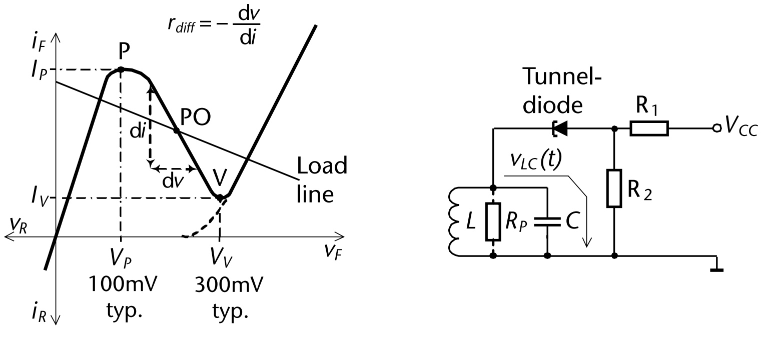
Figure 8.31(a) shows the N-shaped voltage-current characteristic of a tunnel diode (e.g., 1N2927, BD-3, BD-5). If the forward bias voltage is increased, the current through diode increases too until the Esaki peak P is reached (at voltage $V_P$). The next region up to the voltage $V_V$ has a reversed proportionality between voltage and current. Above this threshold, the current rises with the voltage again.

The basic circuit diagram of the oscillator with negative resistance and with a parallel resonant LC circuit is shown in Figure 8.31(b). The tunnel diode is biased by the voltage divider $R_1$ and $R_2$ to operate in the negative resistance region (see the superimposed conductive load line in Figure 8.31(a)).

To obtain maximum output, the quiescent point of operation PO must be placed at the center of the negative-resistance region. Because of this accurate design, the amplitude obtainable from the oscillator with tunnel diode is limited to a fraction of a volt.

The tunnel diodes are very sensitive to temperature changes and overloading and hence are suboptimal for industrial applications in comparison with lambda devices.

A lambda diode is an elementary electronic circuit (Figure 8.32(a)), which combines a complementary pair of JFETs, namely an n-channel type (e.g., MPF102, BF245) and a p-channel type (e.g., 2N3820, 2N5462) into a two-terminal device that exhibits a larger region of differential negative resistance, such as a tunnel diode [21]. The name refers to the shape of the voltage-current characteristic of this device, which resembles the Greek letter lambda ($\lambda$) (Figure 8.32(b)).



(a) (b)

**Figure 8.31** (a) N-shaped voltage-current characteristic of a tunnel diode and the load line for its normally biasing within the negative resistance region, and (b) tunnel diode oscillator (TDO) with parallel resonant LC circuit.

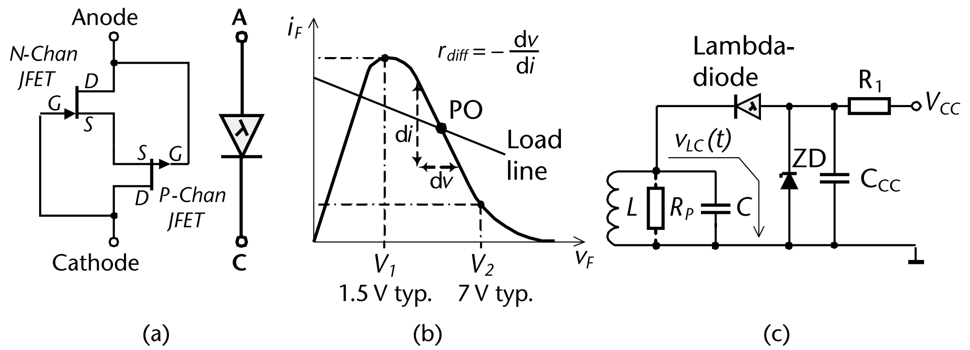


**Figure 8.32** (a) An n-channel JFET (top) and a p-channel JFET combined to form a lambda-diode circuit and the diode symbol, (b) voltage-current characteristic of a lambda diode, and (c) LC oscillator with lambda diode.

Lambda diodes work at a higher voltage than tunnel diodes. The region with differential negative resistance occurs approximately between 1.5 and 7V due to the higher pinch-off voltages of typical JFET devices. A lambda diode therefore cannot replace a tunnel diode directly but is more suitable for industrial applications.

Moreover, in a tunnel diode the current reaches a minimum of about 20% of the peak current before rising again toward higher voltages. The lambda diode current approaches zero as voltage increases, before rising quickly again at a voltage high enough to cause gate-source Zener breakdown in the FETs.

Like the tunnel diode, the negative resistance aspect of the lambda diode lends itself naturally to application in oscillator circuits. Figure 8.32(c) shows a typical oscillation application, whereas the network consisting of $R_1$, Zener-diode, and capacitor $C_1$ provides the biasing of the JFETs.

### 8.3.6.4 Push-Pull Oscillators

A negative resistance characteristic can also be obtained from electronic circuits as, for example, from an astable multivibrator (Figure 8.33(a)) [6, 7]. The oscillating transistors $T_1$ and $T_2$ are cross-coupled by the capacitors $C_1$ and $C_2$.

In absence of the resonant LC circuit, the multivibrator works as a flip-flop circuit and produces square-wave outputs at both collectors, the transistors switching alternately between cutoff and saturation.

The presence of the LC circuit modifies the operation because the inductor provides a low-impedance path between the collectors at low frequencies and the capacitor does the same but at high frequencies, both inhibiting normal multivibrator behavior. Oscillator operation is confined to resonance frequency of the LC circuit at which it has the highest impedance (Section 8.1.2) and the output of the circuit is thus sinusoidal.

A similar schematic with complimentary CMOS-FETs, which can easily be integrated, is shown in Figure 8.33(b). Unfortunately, these transistors have a high 1/f-noise and therefore are not very suitable for oscillator application, except in situations that demand low production costs, which are specific for CMOS technology (see Section 7.3.2).

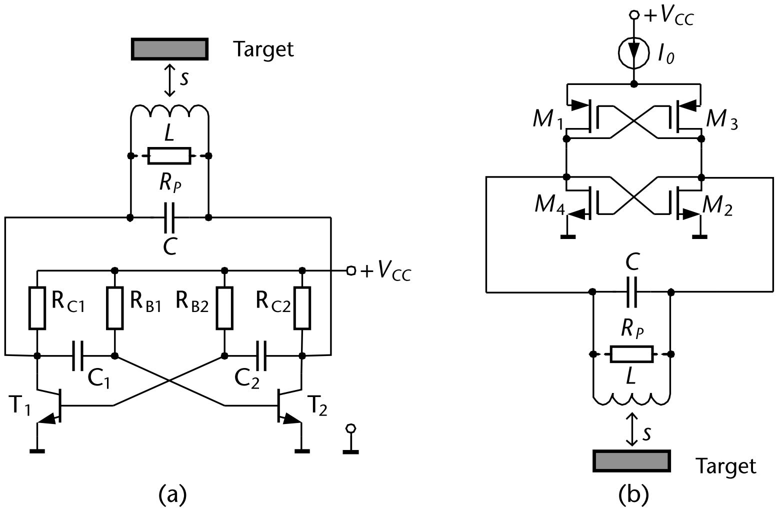


**Figure 8.33** Push-pull oscillators with (a) bipolar transistors, and (b) CMOS transistors.

### 8.3.6.5 Ring-Circuit Oscillator with Additional Negative Resistance Establishment

Section 8.3.6 described two ways to generate negative resistances. Besides the possibility of using specific electronic devices with this property (see Section 8.3.6.3), negative resistance can be achieved within electronic circuits even in schematics with positive feedback [19, 20].

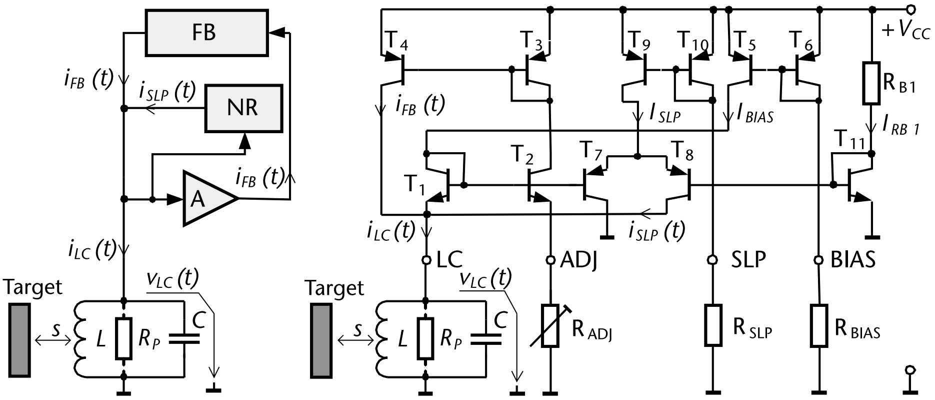
Without going into too much detail, the supplier of the front-end integrated circuit TDA0161, STMicroelectronics, specifies in the corresponding data sheet (see [20] in Chapter 7), that between the connection pins for the resonant LC circuit “the integrated circuit acts like a negative resistance equal to the external resistor $R_1$” connected between the adjusting terminals. Naturally, the solution is intellectual property and is not disclosed.

A similar strategy is the subject of a US patent [22]. That invention has been continuously perfected over the last 20 years and implemented in an ASIC-family (bipolar and CMOS technology) of the Balluff Company (Germany and United States). An important reason for the massive application of these ASICs in inductive sensors serial production is certainly the features of the oscillator, which has the combined functionality of positive feedback and negative resistance.

Other features to be mentioned are the large adjustability and linearity, stability, and compensability for temperature influences, compatibility with a large spectrum of resonant circuits and high EMC immunity, and—last but not least—favorable energetic features. These latter features are only one supply voltage in a large range (2 to 65V) and low current consumptions (below 1 mA). Theoretically, the solution is adequate and advantageous to be integrated in all IC technologies.

The block schematic of the combi solution is shown in Figure 8.34(a). The v-to-i amplifier A and the feedback loop FB form the ring-circuit oscillator and were described in detail in Section 8.3.2. The transistors $T_1$ to $T_6$, carrying the same label as in Figure 8.21(b) are redisplayed in Figure 8.34(b).

The new entity in the block diagram is the negative resistance stage NR, which could have several implementations [22]. One preferential application is the bipolar



(a) (b)

**Figure 8.34** (a) Block diagram of the ring-circuit oscillator with additional negative resistance (NR), and (b) implementation with differential amplifier ($T_7$ and $T_8$).

differential amplifier $T_7$ and $T_8$, biased by an emitter current $I_{SLP}$ that is for example realized with the current mirror $T_9$ to $T_{10}$ and adjusted by the external resistor $R_{SLP}$ [8, 9].

The emitter current $I_{SLP}$ can similarly be expressed as $I_{\text{BIAS}}$:

$$I_{SLP} = \frac{V_{CC} - V_{BE10}}{R_{SLP}} \tag{8.166}$$

The input voltages of the differential amplifier are, on the left-hand side, the sum $v_{LC}(t) + V_{BE1}$, and on the right-hand side the DC-voltage $V_{BE11}$, which can be made to be equal to $V_{BE1}$:

$$V_{BE11} = V_{BE1} \tag{8.167}$$

if the DC current $I_{RB1}$ set by $R_{B1}$ equalizes the main bias current $I_{\text{BIAS}}$. As a result, the differential input voltage of the amplifier is equal to $v_{LC}(t)$.

The output current $i_{SLP}(t)$ of the differential stage is de facto the collector current of the $T_8$ and is directly injected into resonant LC circuit in addition to the previous feedback current $i_{FB}(t)$. Thus, the total current through the LC circuit is:

$$i_{LC}(t) = i_{FB}(t) + I_{\text{BIAS}} + i_{SLP}(t) \tag{8.168}$$

where $i_{FB}(t)$ and its amplitude $I_{fbm}$ are expressed by (8.92) and (8.93), respectively, and $I_{\text{BIAS}}$ by (8.89).

The collector current of the $T_7$ is conducted to GND. The amplifier works as a voltage to current balance. Supposing the condition (8.167) is fulfilled and the oscillator is shut down ($v_{LC}(t) = 0$), the differential amplifier is in a DC equilibrium state, the DC current $I_{SLP}$ is halved, a half flows into the LC circuit without any effect, and the other half flows to GND.

If the oscillator starts to work, the smallest amplitude $V_{lcm}$ (in the range of a few millivolts) is sufficient to periodically swing the differential amplifier ($T_7$ and $T_8$).

During the positive half-wave of $v_{LC}(t)$, the transistor $T_7$ turns off (the higher the amplitude $V_{lcm}$ the smaller the collector current $i_{C7}(t)$). Correspondingly, the voltage of the common emitter connection is held by the transistor $T_8$ on above:

$$V_{E7,8} = V_{BE11} + V_{BE8} \approx 2V_{BE} \tag{8.169}$$

and the surplus of current is taken over by $T_8$.

As a result, the current $i_{SLP}(t)$ shows a positive half-wave over the DC level, namely the level $I_{SLP}/2$. This current is directly injected in the resonant LC circuit.

The negative half-wave of $v_{LC}(t)$ effects a reversed operation. The transistor $T_7$ gradually turns on and provokes a decrease of the voltage $V_{E7,8}$ of the common emitter connection. This voltage dip turns off the transistor $T_8$, whose base voltage is held on a constant voltage $V_{BE11}$.

The output current $i_{SLP}(t)$ follows a negative half-wave below the level $I_{SLP}/2$, their deep looks alike dependant on $V_{lcm}$.

An operating mode with the differential amplifier biased out of the DC equilibrium state, for example, for:

$$V_{BE1} > V_{BE11} \tag{8.170}$$

is often implemented. This gives the possibility to control the degree of influence of the partial output current $i_{SLP}(t)$, by changing the current $I_{\mathrm{BIAS}}$.

After this qualitative analysis, a quantitative analysis has the target to determine the amplitude of the current $i_{SLP}(t)$.

This current, which is the collector current of the transistor $T_8$, is theoretically given by the transfer function of the differential amplifier [10]. If the equilibrium condition (8.167) is fulfilled, the expression of the output current is:

$$i_{SLP}(t) = i_{C8}(t) = \frac{I_{SLP}}{2}\left(1 + \tanh\frac{V_{lcm} \cdot \sin\omega_0 t}{2V_T}\right) \tag{8.171}$$

where $I_{SLP}$ is the bias current of the differential amplifier (see (8.166)) and $V_T$ is the thermal voltage (see (8.103)).

The hyperbolic tangent function $\tanh(x)$ has a quasilinear shape for small arguments ($x \le 1$) and a strong nonlinear aspect for medium and large arguments (Figure 8.35). As a result, the function $\tanh(\sin\omega_0 t)$ in the formula (8.171) and consequently, the current $i_{SLP}(t)$ have a time dependence that is strongly reliant on the ratio between $V_{lcm}$ and $V_T$ (Figure 8.36).

To determine the expression of the effective fundamental wave of the current $i_{SLP}(t)$, which has these vastly different shapes, it would be advised to make a case discrimination based on the oscillation's amplitude $V_{lcm}$. The goal is to convert the expression (8.171) into partial approximation functions, which can easily be expressed by Fourier series.

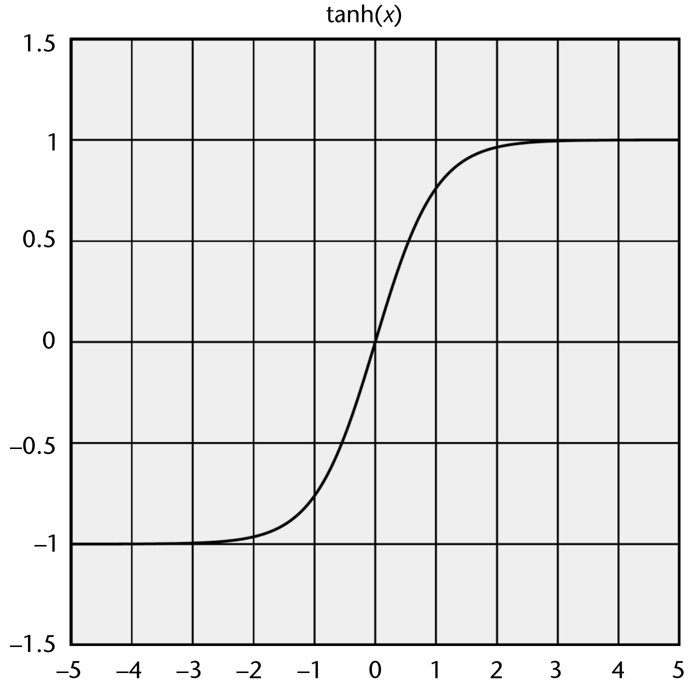


**Figure 8.35** General shape of the hyperbolic tangent function tanh($x$) for $x = -5$ to 5.

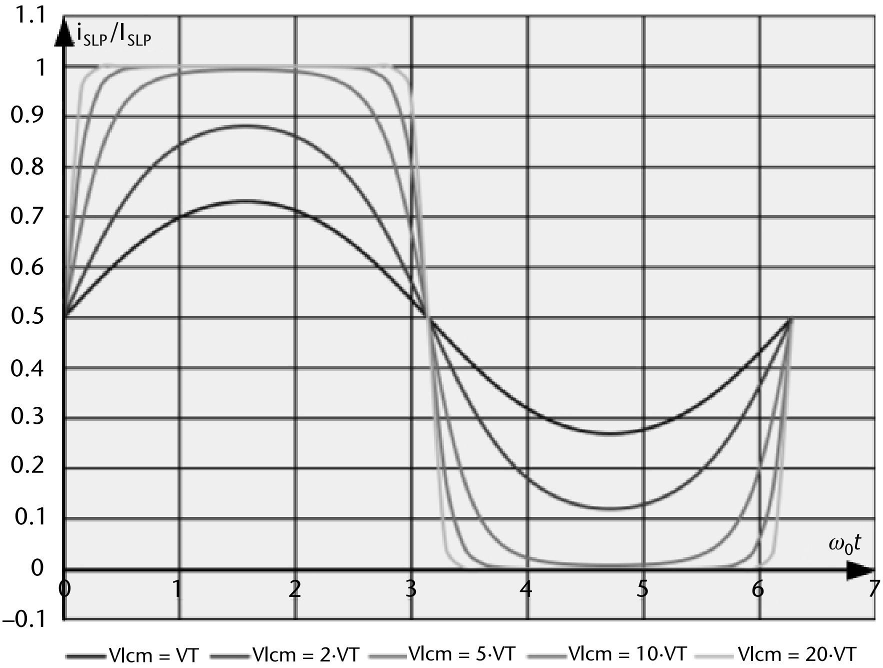


**Figure 8.36** Normalized representation of the current $i_{SLP}(t)$ for different values of the ratio between $V_{lcm}$ and $V_T$.

For small signals, namely for:

$$V_{lcm} \leq 5 \cdot V_T \approx 130 \text{ mV} \tag{8.172}$$

the function tanh($\sin\omega_0 t$) can be replaced by the product of its two parts and the output current results in:

$$i_{SLP}(t) \approx \frac{I_{SLP}}{2}\left[1+\left(\tanh\frac{V_{lcm}}{2V_T}\right)\cdot \sin\omega_0 t\right] \tag{8.173}$$

The amplitude of this approximation sinusoidal wave:

$$I_{slpm} \approx \frac{I_{SLP}}{2}\cdot \tanh\frac{V_{lcm}}{2V_T} \tag{8.174}$$

can be considered unaffected by the small-signal approximation, while the shape of the wave visibly changes (Figure 8.37). The resulting error is shown comparatively in this benchmarking figure.

For medium and high signals ($V_{lcm} > 5 \cdot V_T$), the representation of the current $i_{SLP}(t)$ has aspects that are very close to a trapezoidal shape. The condition (8.172) implies a phase-limit $\alpha_0$ defined as:

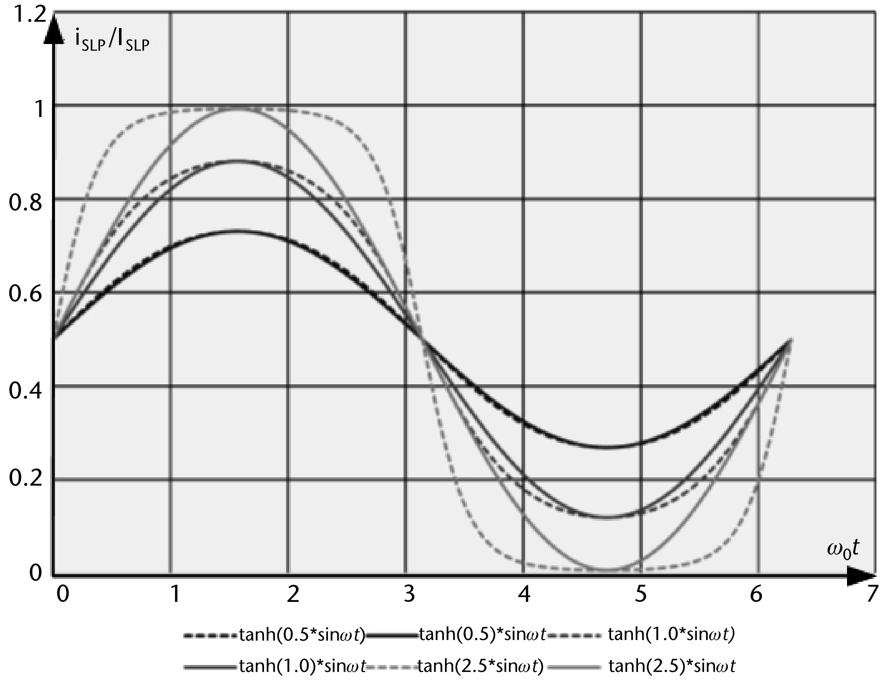


**Figure 8.37** Normalized graphs of the current $i_{SLP}(t)$ according to (8.171) (dashed lines)and to the approximating expression (8.173) (solid lines) for three values of the ratio between $V_{lcm}$ and $V_T$, namely 1, 2, and 5.

$$\frac{V_{lcm}\cdot\sin\omega_0 t}{2V_T} = 2.5 \Rightarrow \alpha_0 = \sin^{-1}\frac{5V_T}{V_{lcm}} \tag{8.175}$$

which can be used to define the trapezoidal approximation function in this case:

$$i_{SLP}(t) \approx \begin{cases} \frac{I_{SLP}}{\alpha_0}\cdot\omega_0 t & \text{for } \omega_0 t \in \left[0,\alpha_0\right] \\ I_{SLP} & \text{for } \omega_0 t \in \left[\alpha_0,\pi-\alpha_0\right] \\ \frac{I_{SLP}}{\alpha_0}\cdot\left(\pi-\omega_0 t\right) & \text{for } \omega_0 t \in \left[\pi-\alpha_0,\pi\right] \end{cases} \tag{8.176}$$

For a realistic signal amplitude of $V_{lcm}$ = 1000 mV, the formula (8.175) gives $\alpha_0 = 8°$. For this reason, the trapezoidal approximation function (8.176) can, at the limit, be replaced by a pure rectangular wave having the expression:

$$i_{SLP}(t) \approx \begin{cases} I_{SLP} & \text{for } \omega_0 t \in [0,\pi] \\ 0 & \text{for } \omega_0 t \in [\pi,2\pi] \end{cases} \tag{8.177}$$

whose Fourier series:

$$i_{SLP}(t) = I_{SLP}\cdot\left[\frac{1}{2}+\frac{2}{\pi}\cdot\left(\sin\omega_0 t+\frac{\sin 3\cdot\omega_0 t}{3}+\frac{\sin 5\cdot\omega_0 t}{5}\right)\right] \tag{8.178}$$

highlights the DC component as well as the odd-order harmonics.

The DC component and the harmonics of the current wave are strongly suppressed (excepting the fundamental wave with the angle frequency $\omega_0$).

This fundamental wave is the only effective component and it has the amplitude:

$$I_{slpm} = \frac{2}{\pi}\cdot I_{SLP} \tag{8.179}$$

and establishes the negative resistance $R_N$:

$$R_N = \frac{V_{lcm}\cdot R_P}{I_{slpm}\cdot R_P - V_{lcm}} = \frac{V_{lcm}\cdot R_P}{\frac{2\cdot I_{SLP}}{\pi}\cdot R_P - V_{lcm}} \tag{8.180}$$

In the formula (8.180) $R_P$ is the well-known equivalent parallel resistance of the lossy parallel LC circuit (see (4.38)).

To determine the influence of the negative resistance on the dynamic behavior of the oscillator let us consider this negative resistance $R_N$ connected in parallel with the resonant LC circuit, which is now represented by its physical equivalent circuit (see Figure 4.6(b)).

The total admittance, expressed in the Laplace transform domain, is the sum of the three partial admittances:

$$\underline{Y}(s) = \underline{Y}_L(s) + \underline{Y}_C(s) + \underline{Y}_{RN}(s) = \frac{1}{R_s + sL} + sC + \frac{1}{R_N} \tag{8.181}$$

where the series equivalent resistor is expressed in (5.35).

Correspondingly, the total impedance results in:

$$\underline{Z}(s) = \frac{1}{\underline{Y}(s)} = \frac{\left(sL + R_s\right) \cdot R_N}{s^2 LCR_N + s\left(CR_s R_N + L\right) + \left(R_s + R_N\right)} \tag{8.182}$$

The expression of the oscillator time constant can be computed by means of the oscillator time response to a current pulse in form of the Dirac delta function $\delta(t)$ [5], also known as the unit impulse. It is an elementary pulse having a unitary height and width tending to zero. Its Laplace transform is equal to one and thus the Laplace transform of the circuit response is:

$$\underline{V}_{LC}(s) = \underline{Z}(s) \cdot \mathcal{L}\{\delta(t)\} = \frac{\left(sL + R_s\right) \cdot R_N}{s^2 LCR_N + s\left(CR_s R_N + L\right) + \left(R_s + R_N\right)} \tag{8.183}$$

Returning from the complex domain to the time domain, the time response to this impulse is given by the inverse Laplace transform:

$$v_{LC}(t) = \mathcal{L}^{-1}\{\underline{V}_{LC}(s)\} = \mathcal{L}^{-1}\left\{\frac{\left(sL + R_s\right) \cdot R_N}{s^2 LCR_N + s\left(CR_s R_N + L\right) + \left(R_s + R_N\right)}\right\}. \tag{8.184}$$

The quadratic denominator of the characteristic polynomial in (8.184) can be factorized and the general expression of the time response results in:

$$v_{LC}(t) = \mathcal{L}^{-1}\left\{\frac{s}{\left(s + s_1\right) \cdot \left(s + s_2\right)}\right\} = \frac{1}{s_1 - s_2}\left(s_1 e^{-s_1 t} - s_2 e^{-s_2 t}\right) \tag{8.185}$$

where $s_1$ and $s_2$ are the system poles (roots of the denominator).

In the classical formula, to calculate the roots of a quadratic equation:

$$s_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{8.186}$$

with the coefficients concretely expressed for the present system by:

$$a = LCR_N,\ b = \left(CR_S R_N + L\right) \text{ and } c = R_S + R_N \tag{8.187}$$

the value of the determinant is negligible for real oscillators for inductive sensors and so that the roots are real quantities:

$$s_{1,2} \approx \frac{-b}{2a} = \frac{-\left(CR_S R_N + L\right)}{2 \cdot LCR_N} \tag{8.188}$$

For this pair of poles placed on the right-hand side of the complex plane (see Figure 8.11), the response to an impulse has a time constant:

$$\tau = -\frac{1}{s_{1,2}} \approx -\frac{2a}{-b} = \frac{2 \cdot LCR_N}{CR_S R_N + L} = \frac{2LC}{CR_S + \dfrac{L}{R_N}} \tag{8.189}$$

If $R_N \to \infty$, the equation reduces to the expression:

$$\tau = \frac{2L}{R_S} \tag{8.190}$$

The negative resistance has a beneficial influence, reducing the time constant. Considering (8.180), substituting this into (8.189), simplifying and reducing it, finally results in a time constant:

$$\tau = \pi \cdot \frac{C \cdot V_{lcm}}{I_{SLP}} \tag{8.191}$$

At the end, the oscillator current consumption can immediately be expressed referring to the formula (8.105). In opposition to the oscillator in Figure 8.21(b), the improved oscillator in Figure 8.34(b) has a higher current consumption due to the additional stage with $T_7$ ... $T_{10}$ and the biasing diode realized with $T_{11}$.

Consequently, the current consumption is given by:

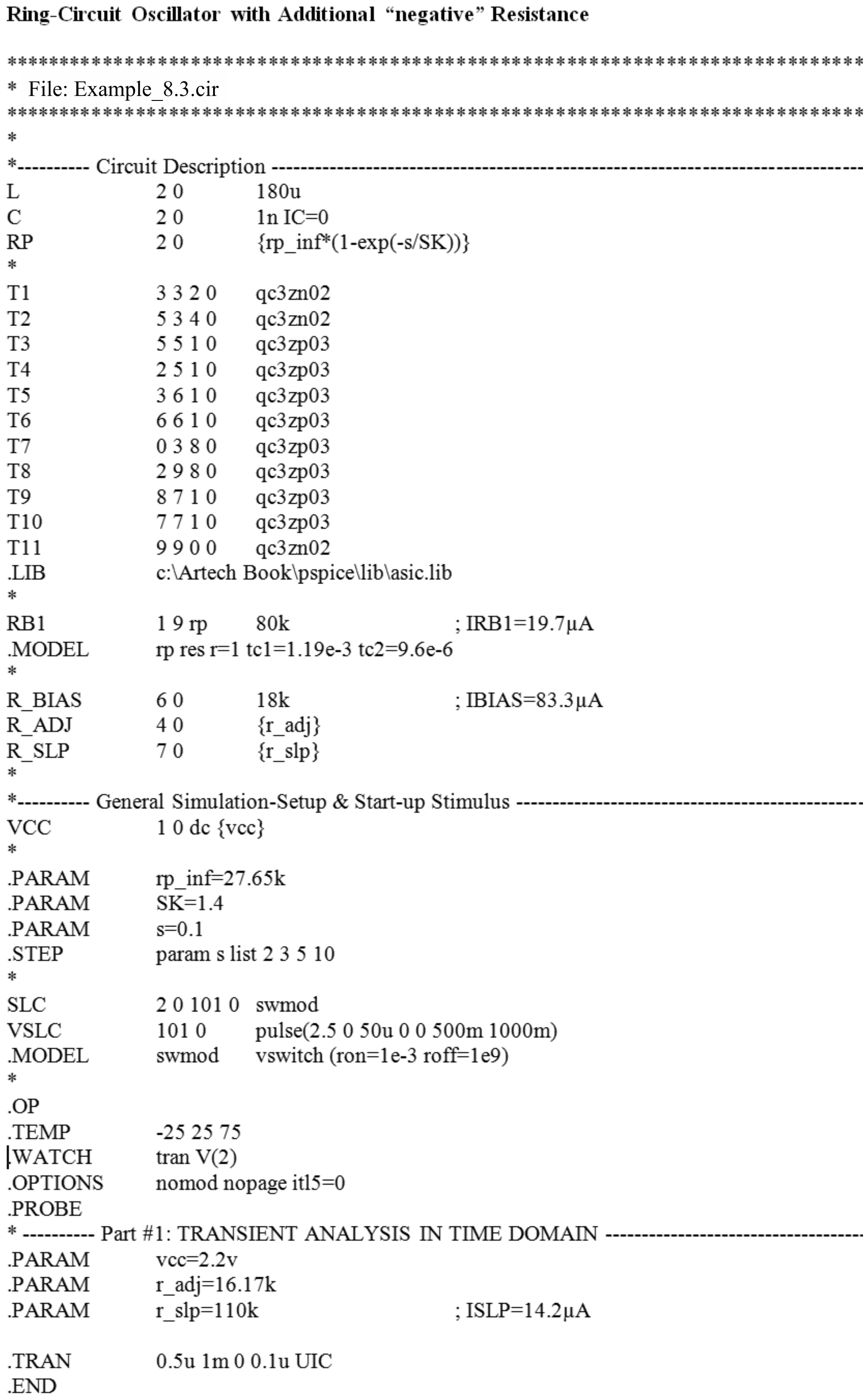
$$I_{CC} = 2 \cdot I_{BIAS} + 2 \cdot \bar{I}_{FB} + 2 \cdot I_{SLP} + I_{RB1} \tag{8.192}$$

where $\bar{I}_{FB}$ is the average values of the current $i_{FB}(t)$ (see (8.106)), $I_{BIAS}$ and $I_{SLP}$ are the currents through $R_{BIAS}$ and $R_{SLP}$, respectively, and $I_{RB1}$ is the current flowing through $R_{B1}$ and $T_{11}$.

*Example 8.3*

Using the classical circuit simulation software SPICE [23], the oscillator in Figure 8.34(b) was simulated under the following working conditions (see the circuit file *.cir* in Figure 8.38):

- Lossy parallel resonant LC circuit: Losses are modeled with the expression (8.82). Quantities $R_{P\infty}$, $s_0$, and $SK$ are known, the distance $s$ is a parameter.
- Low supply voltage (2.2V), specific for up-to-date low-voltage systems.
- Three ambient temperatures: –25°C, 25°C, and 75°C.
- DC biasing: $I_{BIAS} \approx 84\ \mu A$, $I_{SLP} = 14.2\ \mu A$, typ.
- $R_{B1}$ is an integrated base resistor (see the step "base processing" in Figure 6.9) with typical temperature coefficients [10]: linear *tc*1 and quadratic *tc*2.



**Ring-Circuit Oscillator with Additional "negative" Resistance**

```
**************************************************************************
* File: Example_8.3.cir
**************************************************************************
*
*---------- Circuit Description ------------------------------------------------------------------------------
L              2 0        180u
C              2 0        1n IC=0
RP             2 0        {rp_inf*(1-exp(-s/SK))}
*
T1             3 3 2 0    qc3zn02
T2             5 3 4 0    qc3zn02
T3             5 5 1 0    qc3zp03
T4             2 5 1 0    qc3zp03
T5             3 6 1 0    qc3zp03
T6             6 6 1 0    qc3zp03
T7             0 3 8 0    qc3zp03
T8             2 9 8 0    qc3zp03
T9             8 7 1 0    qc3zp03
T10            7 7 1 0    qc3zp03
T11            9 9 0 0    qc3zn02
.LIB           c:\Artech Book\pspice\lib\asic.lib
*
RB1            1 9 rp     80k                         ; IRB1=19.7µA
.MODEL         rp res r=1 tc1=1.19e-3 tc2=9.6e-6
*
R_BIAS         6 0        18k                         ; IBIAS=83.3µA
R_ADJ          4 0        {r_adj}
R_SLP          7 0        {r_slp}
*
*---------- General Simulation-Setup & Start-up Stimulus ------------------------------------------------
VCC            1 0 dc {vcc}
*
.PARAM         rp_inf=27.65k
.PARAM         SK=1.4
.PARAM         s=0.1
.STEP          param s list 2 3 5 10
*
SLC            2 0 101 0  swmod
VSLC           101 0      pulse(2.5 0 50u 0 0 500m 1000m)
.MODEL         swmod      vswitch (ron=1e-3 roff=1e9)
*
.OP
.TEMP          -25 25 75
|WATCH         tran V(2)
.OPTIONS       nomod nopage itl5=0
.PROBE
* ---------- Part #1: TRANSIENT ANALYSIS IN TIME DOMAIN ------------------------------------
.PARAM         vcc=2.2v
.PARAM         r_adj=16.17k
.PARAM         r_slp=110k                             ; ISLP=14.2µA

.TRAN          0.5u 1m 0 0.1u UIC
.END
```

**Figure 8.38** An example of PSpice circuit file that uses a transient analysis to evaluate the behavior of the oscillator in Figure 8.34(b) and the phase plane portraits (see Section 8.2).

- The external resistors $R_{ADJ}$, $R_{BIAS}$, and $R_{SLP}$ are considered without temperature drift, which is fulfilled by the up-to-date real resistors.
- Typical PSpice models for integrated bipolar transistors.

Besides the classical operating point OP computation and monitoring features (*WATCH* and *PROBE*), the main task is the time-domain analysis provided by the command *.TRAN*. The limited space in this book does not allow to present too many results and plots. A restricted number of oscillator specific plots are still exposed.

Figure 8.39 shows gained records of the planes of the states of the capacitor and coil currents versus node voltage V(2), which is the voltage $v_{LC}(t)$, during the startup interval (from initial time $t = 0$ up to the instant of achievement of oscillator steady state). The regular convergent oval-shaped plots denote an oscillator operating mode with harmonic oscillations (see Figure 8.7).

Figure 8.40 shows the planes of the states of the currents $i_{C7}$ and $i_{C8}$ versus the same node voltage and under similar conditions. The difference between the currents $I_{\mathrm{BIAS}}$ and $I_{RB1} \approx 20\ \mu\mathrm{A}$ causes a shifting of the crossing point. The limiting level of the currents imposed by the current $I_{SLP} \approx 14\ \mu\mathrm{A}$ is immediately visible.

To conclude, the features of the pervious ring-circuit oscillators in Figure 8.21(b) are also available for the ring-circuit oscillator with additional negative resistance (see the conclusions in Section 8.3.2).

The great advantage of the second embodiment in Figure 8.34(b) lies in the possibility to control the oscillator characteristic curve: amplitude $V_{lcm}$ versus distance $s$ between coil and target. Both the slope and measuring range but also the linearity of this curve can be adjusted corresponding to the application specification.

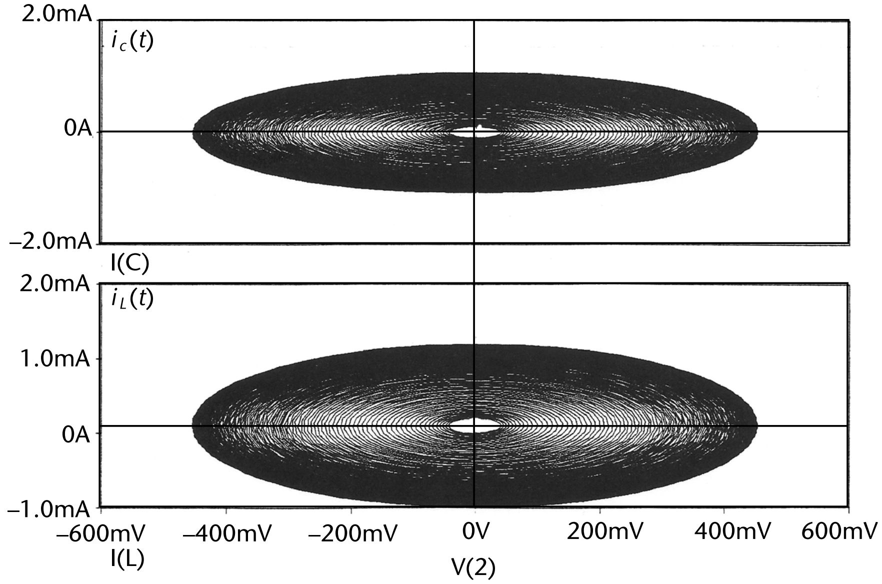


**Figure 8.39** PSpice simulation results: Planes of the states of the capacitor current $i_C(t)$ and coil current $i_L(t)$ versus node voltage V(2); that is, the voltage $v_{LC}(t)$ (compare with Figure 8.7). The small offset of $i_L(t)$ at $v_{LC}(t) = 0$ measures the bias currents through the coil.

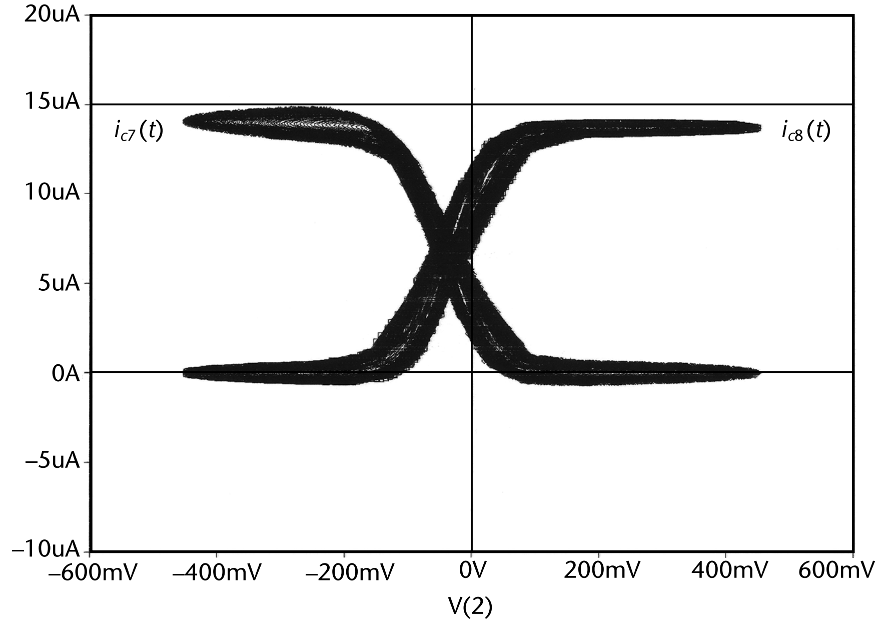


**Figure 8.40** PSpice simulation results: Planes of the states of the currents $i_{C7}$ and $i_{C8}$ versus the node voltage V(2); that is, the voltage $v_{LC}(t)$.

Moreover, the additional circuitry makes an efficient control of the oscillator temperature behavior possible so as to compensate intrinsic temperature drifts of some sensor components such as—first of all—the coil.

The circuitry is predestined to be integrated including the resistor $R_{B1}$ as integrated base resistor [10]. $R_{ADJ}$, $R_{BIAS}$, and $R_{SLP}$ remain external adjusting and/or setup sensor components.

## 8.4 Function Generators

A chapter that examines oscillators for inductive sensors cannot be completed without describing, even in summary form, the second class of oscillators implemented in inductive sensors. As already mentioned in the introduction of this chapter, they do not operate under electromagnetic influence of the target and thus, their working parameters (amplitude, frequency and/or phase of the provided signal) are theoretically constant. Their tasks are to supply with HF signals various circuitry constituting inductive sensing elements or to clock electronic stages in the sensor evaluation electronics.

These types of oscillators are not strictly inductive sensor-specific and can be found in several volumes of electronic engineering [7, 17].

The presentation of function generators primarily focuses on the output signal shape (variation in time) and secondarily on the frequency and amplitude of this signal.

### 8.4.1 Relaxation Oscillators

Relaxation oscillators generally deliver triangle and rectangular (square wave) signals having as key part an RC oscillating core but also an LC circuit.

To completely evaluate the operation of an oscillator, it is necessary to write and solve the system of specific differential equations describing the circuit, which is generally a difficult task. Academic publications disclose classical general differential equations, which should describe the oscillator behavior. Such a modeling is offered by the Van der Pol equation, which is valid for traditional but also for relaxation oscillators [2]:

$$\ddot{x} + \varepsilon\left(x^2 - 1\right)\dot{x}^2 + \omega_0 x = 0 \tag{8.193}$$

where $x$ is the system output, $\ddot{x}$ the second derivative of $x$ with respect to time, $\omega_0$ is the angular frequency of the system, and the parameter $\varepsilon$ defines the damping degree.

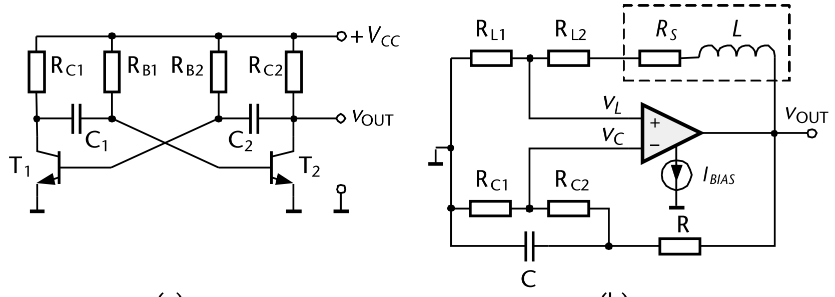
Thus, the oscillator behavior depends on the value of $\varepsilon$. For $\varepsilon = 0$, (8.193) reduces to (8.2) and describes a harmonic oscillator. The oscillators described in the previous sections are characterized by a low $\varepsilon$ value and the nonlinearity imposed by the middle term in (8.193) is negligible.

Should the $\varepsilon$ value increase, the output signal shape fluently changes from a sinusoid form to a square wave. For $\varepsilon >> 5$, the signal shape is rectangular with low distortion factor [2].

The earliest relaxation oscillator is the classical astable multivibrator (Figure 8.41(a)). The circuit is frequently present in technical publications [2, 6], and its operation was concisely described in Section 8.3.6.4. In general, the circuit has a symmetrical structure ($C_1 = C_2 = C$ and $R_{B1} = R_{B2} = R$) and the output signal $v_{\mathrm{OUT}}(t)$ is a symmetrical square wave with the levels equal to GND and $V_{CC}$ and with the repeat frequency:

$$f = \frac{1}{T} = \frac{1}{2\mathrm{RC} \cdot \ln 2} \approx \frac{1}{1.386 \cdot \mathrm{RC}} \tag{8.194}$$

Apart from various RC-based multivibrators, professional papers present LC-based versions of relaxation oscillators. Section 5.5 describes, for example, an inductive proximity microsensor, which is composed of a planar sensing coil on



(a) (b)

**Figure 8.41** Relaxation oscillators: RC flip-flop with bipolar transistors (a) and LC version (b).

a substrate and integrated evaluation electronics, containing an oscillator with a single-ended op-amp (Figure 8.41(b)). This is an LC relaxation oscillator with an op-amp in differential configuration, which generates a squared AC output signal.

The oscillator coil is represented in Figure 8.41(b) by its equivalent circuit, $L$ in series with $R_S$ (see Figure 4.6(b)), and is inserted in the positive feedback path. The negative feedback path contains the oscillating capacitor $C$. The steady state is achieved for a suitable bias current $I_{\mathrm{BIAS}}$, which controls the amplification of the op-amp. The output signal frequency is given by the well-known Thomson's formula:

$$f = \frac{1}{2 \cdot \pi\sqrt{LC}} \tag{8.195}$$

This circuit can be used either as a function generator or as an oscillator for the proximity metal object detection. The presence of the inductive part gives a better frequency stability compared to the RC multivibrator, whose working frequency is very dependent on the instable BE and CE saturation voltages of the transistors.

A second basic circuit that belongs to the relaxation oscillators is the astable blocking oscillator with reduced electronics but with an additional transformer [6]. Due to this expensive part, this version is no longer up to date.

### 8.4.2 Self-Oscillating Function Generators

Self-oscillating generators can naturally be realized with discrete electronic components; nevertheless, the availability today of rail-to-rail op-amps makes sense to use these amps for such implementations.

An astable multivibrator with an op-amp is illustrated in Figure 8.42(a). Initially, the capacitor C will be charged through the resistor R. This happens until the voltage at the inverting input of the op-amp exceeds the voltage at the noninverting input, which is set by the voltage divider $R_1$ and $R_2$ that is momentarily supplied with the HIGH level of the op-amp output voltage.

In this moment, due to its high gain, the op-amp switches on LOW and the discharging of the capacitor through R is started. This occurs until the capacitor voltage falls below the new voltage at the noninverting input, which is determined by the same voltage divider but right now supplied with the LOW level of the op-amp output voltage.

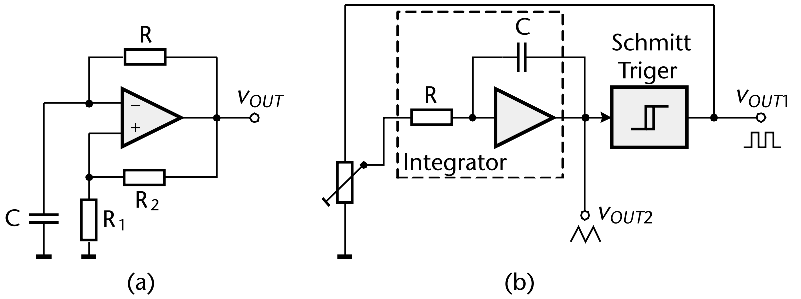


**Figure 8.42** Astable multivibrator (a) and triangle signal generator (b) with op-amps.

The op-amp switches again, the output goes on HIGH, and a new charging cycle starts. The process theoretically recurs as long as the schematic is supplied. For a symmetrical supplying, the charging/discharging times are equal and do not depend on the supply voltages [2]:

$$t_{CH} = t_{DCH} = RC \cdot \ln \frac{2R_1}{R_1 + R_2} \tag{8.196}$$

This symmetric duty cycle:

$$D_C \overset{\text{def}}{=} \frac{t_{CH}}{t_{CH} + t_{DCH}} = 0.5 \tag{8.197}$$

can be modified by changing the levels of the supply voltages.

A self-oscillating triangle signal generator is shown in Figure 8.42(b). The schematic is suitable as core piece of universal generators for various signals. In parallel to the triangle signal, the circuitry also delivers a rectangular signal.

The schematic consists of an integrator with op-amp and a Schmitt trigger (see Section 9.6), which can be realized with a second op-amp (e.g., $OA_3$ in Figure 7.15). The stages are connected into a total feedback loop. The rectangular output signal, available at the trigger output, is fed back to the integrator input. The trigger input is the triangle signal, which is generated by the integrator by means of the rectangular output signal. Thus, a positive output signal value causes a linear falling triangle signal. The loop gain is also high (op-amps have high gain values), thus the conditions of the Barkhausen criterion are fulfilled and the schematic continuously oscillates.

The amplitude of the intermediate triangle voltage is adjusted by the voltage divider present at the integrator input. In this way, the divider controls the working frequency, which can be expressed by [2]:

$$f \approx \frac{N \cdot V_{\text{OUT}}}{4RC \cdot V_{TH}} \tag{8.198}$$

where $V_{\text{OUT}}$ is the level (positive or negative) of the symmetrical trigger output signal, $V_{TH}$ is its threshold level (positive or negative), and $N$ is the dividing factor of the input potentiometer (usual values between 0.1 and 1).

Both square wave and sawtooth signals can also be gained by means of the integrated functions generator NE566 (manufacturer NXP company). The device is a voltage-controlled oscillator of exceptional linearity with buffered square wave and triangle wave outputs. The frequency of oscillations is determined by external resistor and capacitor and by a voltage applied to the control terminal. The circuit simultaneously provides square wave and triangle wave signals at frequencies up to 1 MHz. The oscillator can be programmed over a ten-to-one frequency range by proper selection of the external resistance. Optionally, the signals can be modulated by the control voltage $V_M$, with high linearity.

A typical connection diagram is shown in Figure 8.43. The control terminal (Pin 5) must be biased externally with a voltage $V_5$, which is provided by a voltage

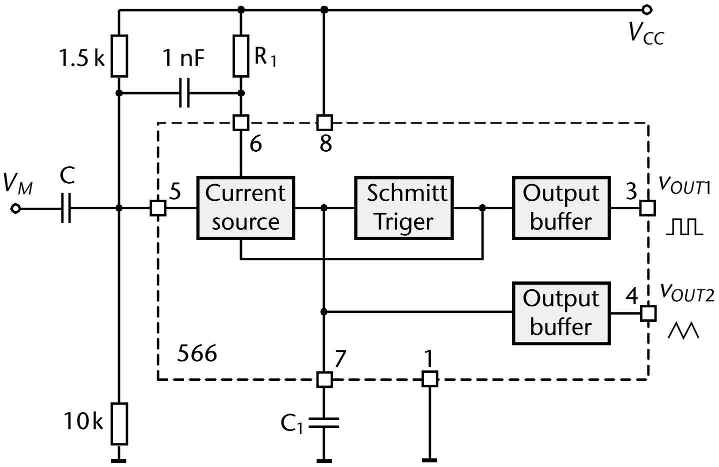


**Figure 8.43** Triangle and rectangular signal generator with the integrated circuit 566.

divider formed with 1.5 kΩ and 10 kΩ. The modulating signal $V_M$ (if any) is then AC-coupled with the capacitor C. The modulating signal could be directly coupled as well, if the appropriate DC bias voltage is applied to the control terminal.

The frequency is given approximately by [24]:

$$f \approx \frac{2 \cdot \left(V_{CC} - V_M\right)}{R_1 C_1 V_{CC}} \tag{8.199}$$

where $R_1$ should be in the range between 2 and 20 kΩ.

A small capacitor (typically 1 nF) should be connected between Pins 5 and 6 to eliminate possible oscillation in the controlled current source.

A specific application of the function generators that provide both signal shapes is the generation of a rectangular signal pair: the in-phase square wave directly obtained from the corresponding output, and the quadrature square wave gained by an inverting amplifier that is supplied with the triangle signal.

### 8.4.3 Timer-Chip NE555

Finally, a volume section called function generators cannot be completed without the presentation of the classical and brilliant integrated timer circuit NE555.

In the entire electronics history, over the years 555 has been the absolutely most successful device because it has innumerable applications and it is the best-selling integrated circuit (one billion sold parts as of 2003). With this marketed amount, 555 exceeded its older competitor, the legendary op-amp $\mu$A741.

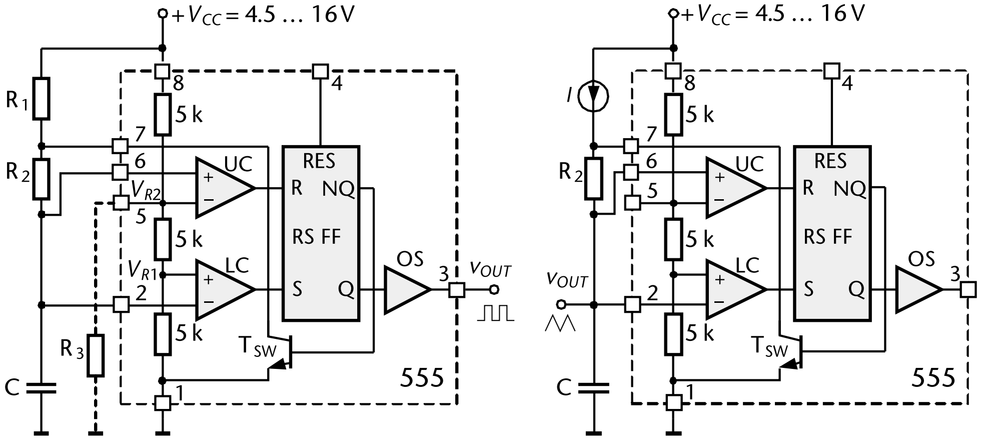
Hans R. Camenzind, also know as “Mister 555” invented the circuit in 1971 and his Signetics Company in Sunnyvale/Santa Clara (later Philips Semiconductors, today NXP Semiconductors) immediately started batch production. After the product launch in 1972, the circuit was taken over by large groups like Texas Instruments (TLC555) and National (LMC555). Eighteen years later, a CMOS version was established. Versions of the device were made or are still made after 40 years by dozens of major semiconductors vendors, including Texas Instruments, Intersil-Renesas, Maxim Integrated, Exar, Fairchild Semiconductor, STMicroelectronics, and NXP. The CMOS version has been improved over the years, but the bipolar design has never been changed (see the schematic in [26]).

A large number (more than 20) of technical books deal with 555 operation and applications as (1) a timer, signal generator, and device for optoelectronic circuits, (2) a part for science and communication circuits and projects, and (3) a device for electronic sensor circuits and projects. Some significant applications for the subject matter of our volume are:

- Monostable and astable multivibrator;
- Comparator, Schmitt trigger (see Section 9.6);
- Pulse detector, pulse modulator;
- Frequency divider, timing generator;
- Square wave generator;
- Sawtooth and ramp signal generator, etc.

The chip architecture is ingenious and straightforward [25, 26] and consists of the following parts (Figure 8.44(a)):

- Triple voltage divider (3 × R = 5 kΩ);
- Two comparators, UC and LC;
- One RS flip-flop with additional reset input RES;
- One switching open-collector transistor $T_{SW}$;
- Output stage OS.



(a) (b)

Pin configuration

1 GND
2 Trigger
3 Output
4 Reset
5 Control voltage
6 Threshold
7 Discharge
8 VCC

**Figure 8.44** A stable multivibrator (a) and triangle signal generator (b) with integrated circuit 555.

The schematic in Figure 8.44(a) shows an astable multivibrator application. Without external components, the 555 works as comparator with hysteresis. If the TRIGGER input voltage (pin 2) goes below the reference voltage $V_{R1} = V_{CC}/3$, the comparator LC sets the flip-flop and the signal OUTPUT (pin 3) goes HIGH. If the THRESHOLD input voltage (pin 6) goes above the reference voltage $V_{R2} = 2 \cdot V_{CC}/3$, the comparator UC resets the flip-flop and the signal OUTPUT (pin 3) goes LOW.

The external periphery changes the operation into an astable multivibrator mode.

When the supply voltage is applied, the capacitor C charges through the series network $R_1$ and $R_2$ until its voltage exceeds the reference voltage $V_{R2}$.

In this moment, the comparator UC toggles the flip-flop. Its output Q switches on LOW and the main OUTPUT (pin 3) goes LOW. The inverted flip-flop output NQ turns on the switching transistor $T_{SW}$ and a discharging of the capacitor through $R_2$ and $T_{SW}$ is started. This occurs until the capacitor voltage falls below the reference voltage $V_{R1}$.

Now the flip-flop is again toggled by the comparator LC and the OUTPUT (pin 3) switches on HIGH. The inverted flip-flop output NQ turns off the switching transistor $T_{SW}$ and a new charging cycle starts. It continues until the reference voltage $V_{R2}$ is achieved again.

The charging time is expressed by:

$$t_{CH} = (R_1 + R_2)C \cdot \ln\frac{V_{CC} - V_{R1}}{V_{CC} - V_{R2}} = (R_1 + R_2)C \cdot \ln 2 \tag{8.200}$$

and the discharging time is given by:

$$t_{DCH} = R_2 C \cdot \ln\frac{V_{R2}}{V_{R1}} = R_2 C \cdot \ln 2 \tag{8.201}$$

Consequently, the frequency of the output signal is:

$$f = \frac{1}{t_{CH} + t_{DCH}} = \frac{1}{T} \approx \frac{1}{0.693 \cdot (R_1 + 2R_2)C} \tag{8.202}$$

and its duty cycle results in:

$$D_C \overset{\text{def}}{=} \frac{t_{CH}}{T} = \frac{0.693 \cdot (R_1 + R_2)C}{0.693 \cdot (R_1 + 2R_2)C} = \frac{R_1 + R_2}{R_1 + 2R_2} \tag{8.203}$$

Note that the signal parameters are independent of the supply voltage and its variations. The duty cycle can be modified by means of an additional external resistor $R_3$ connected between the input CONTROL (pin 5) and GND, which changes the reference voltages $V_{R1}$ and $V_{R2}$. To get symmetrical square wave with $D_C = 0.5$, this resistor should be $R_3 = 2^{1/2} \cdot R \approx 7.071$ kΩ.

The major changes in Figure 8.44(a) to get a triangle signal generator are to replace the resistor $R_1$ by a DC current source (e.g., electronic circuit as shown in Figure 8.21(b) for the bias current $I_{\text{BIAS}}$ or a current regulator diode CRD as in

Figure 7.16(a), etc.) and to use the capacitor voltage (directly or buffered) as output voltage $v_{OUT}$ (Figure 8.44(b)).

The DC current $I$ linearly charges the capacitor $C$ between $V_{R1}$ and $V_{R2}$ in time interval:

$$t_{CH} = \frac{V_{CC}}{3 \cdot I} \cdot C \tag{8.204}$$

and generates the rising ramp of the sawtooth signal.

The discharging of the capacitor $C$ between $V_{R2}$ and $V_{R1}$ occurs again through $R_2$ and $T_{SW}$, and has the same duration as in (8.201):

$$t_{DCH} = R_2 C \cdot \ln \frac{V_{R2}}{V_{R1}} = R_2 C \cdot \ln 2 \tag{8.205}$$

The times can be adjusted: both times by $C$ and independently each other by the current $I$ and $R_2$, respectively. However, the lowest value of $R_2$ is limited by the highest current allowed through the transistor $T_{SW}$ (1 mA, typ.).

## 8.5 Digitally Synthesized Function Generators

In contrast to the above described generators, the digital function generators are characterized by exploiting of a numerically depicted variation of the output signal, which is finally provided at the generator output(s) as digital and/or analog signal.

The block diagram in Figure 8.45 shows the basic structure of a digital function generator. From a systemic point of view, it is a function convertor that converts the clock signal produced by an internal clock oscillator into output functions.

The frequency divider FDIV gradually and also continuously changes the frequency of the clock generator CLKG giving the final frequency of the generated signals. The next counter stage COUNT quantizes one period of the signal oscillation. The counter runs up and step by step addresses the locations of a memory unit MEM, in which the curve progression of the output signal is stored, namely

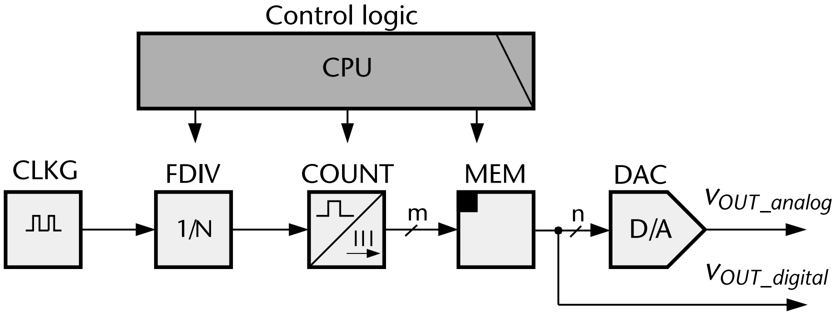


**Figure 8.45** Block schematic of a digital functions generator.

the amplitude versus the time. This time discrete function is directly provided as digital signal. To get an analog output signal, the digital pattern is converted into an analog signal by a DAC.

The nonvolatile memory can store one or more signal patterns for different output functions (sinusoidal, square, triangle, etc.). Up-to-date generators contain full-power CPUs that control large memory capacities. These control units can compute the mathematical description of the signal patterns or can write and as needed read the digitized signal patterns (LUT procedure).

Without a doubt, the method is the procedure of the modern times of digitalization. Note that the disadvantage consisting of occurring quantization noises can be minimized by increasing the resolution in time and in function value. That makes a corresponding increase of the conversion speed necessary, which is also easier to master at the present time.

## References

[1] Webster, J. G. (Editor-in-Chief), *The Measurement, Instrumentation and Sensors Handbook*, Boca Raton, FL: CRC Press/IEEE Press, 1999.

[2] Kurz, G., *Oszillatoren*, Heidelberg, Germany: Hüthig-Verlag 1994.

[3] Andronov, A. A., A. A. Vitt, and S. E. Khaikin, *Theory of Oscillators*, New York: Dover Publications, Inc., 1966.

[4] Dorf, R. C. (Editor-in-Chief), *The Electrical Engineering Handbook*, Boca Raton, FL: CRC Press/IEEE Press, 1997.

[5] Nearing, J., *Mathematical Tools for Physics*, ebook, Miami: Dover Publications, 2010.

[6] Amos, S., and M. James, *Principles of Transistor Circuits*, Ninth Edition, ebook, Oxford, UK: Newnes, 2003.

[7] Tietze, U., C.Schenk, and E. Gamm, *Halbleiter-Schaltungstechnik*, 15th Edition, Berlin, Germany: Spinger/Vieweg-Verlag, 2016.

[8] Jagiella, M., and S. Fericean, "Miniaturized Inductive Sensors for Industrial Applications," *Proceedings of the First IEEE Conference on Sensors*, Orlando, FL, 2002.

[9] Fericean, S. and Droxler, R., "New Noncontacting Inductive Analog Proximity and Inductive Linear Displacement Sensors for Industrial Automation," *Sensors Journal*, Vol. 7, No. 11, 2007.

[10] Gray, P. R., P. J. Hurst, S. H. Lewis, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, Fifth Edition, New York: John Wiley & Sons, 2009.

[11] Fujimori, N. et al., Omron Corporation (Japan), Proximity Sensor, DE Patent 10 2005 014492B4, December 31, 2015.

[12] Niwa, M.,et al., Panasonic Corporation (Japan), Proximity Sensor, U.S. Patent 8,432,169, issued April 30, 2013.

[13] Heimlicher, P., Proximity Sensor, U.S. Patent Application 2014/0117979, May 1, 2014.

[14] Boon, Cornelius Alexander Maria / Philips (Netherlands), Oscillator Circuit, EP Patent 0 261 714, July 31, 1997.

[15] Eissler, W., et al., *Praktischer Einsatz von berührungslos arbeitenden Sensoren*, 2nd Edition, Ehningen, Germany: Expert-Verlag, 1996.

[16] Riccardo Condorelli, Daniele Mangano / STMicroelectronics (IT), *Method of Interfacing an LC Sensor and Related System*, US-Patent 9,897,630, Feb. 20, 2018.

[17] Christiansen, D., and A. Charles, *Standard Handbook of Electronic Engineering*, Fifth Edition, New York: McGraw-Hill Professional, 2005.

[18] Reitsma, G., Resonant Impedance Sensing based on Controlled Negative Impedance, U.S. Patent 9,088,261, issued July 21, 2015.

[19] Reitsma, G., Resonant Impedance Sensing with a Negative Impedance Control Loop Implemented with Synchronized Class D and Output Comparators, U.S. Patent 9,638,763, issued May 2, 2017.

[20] M. M. Altarawneh, et al., "Proximity Detector Circuits: An Alternative to Tunnel Diode Oscillators for Contactless Measurements in Pulsed Magnetic Environments," *Review of Scientific Instruments*, Vol. 80, 2009.

[21] http://www.softrockradio.org/items-of-interest/lambda-diodes, How to Create Lambda Diodes and Design Circuits with Them.

[22] Fericean, S., M. Friedrich, and E. Gass, *Inductive Sensor Responsive to the Distance to a Conductive or Magnetizable Object*, U.S. Patent 5,504,425, issued April 2, 1996.

[23] Vladimirescu, A., *The SPICE Book*, New York: John Wiley & Sons, 1994.

[24] http://www.alldatasheet.com/datasheet-pdf/pdf/17982/PHILIPS/NE566.html, NE566 Datasheet, NXP Semiconductors, 2016.

[25] Camenzind, H., *Designing Analog Chips*, ebook, Book Surge Publishing, 2005.

[26] www.ti.com/lit/ds/symlink/lm555.pdf, LM555 Timer (Rev. D), Texas Instruments, Jan-uary 2015.

CHAPTER 9

# Inductive Sensors: Signal Processing and Conditioning

Chapter 8 exclusively treated the core piece of the evaluation electronics of ISs, the oscillator. This chapter is devoted to the remaining stages of this evaluation electronics (see the generic block diagram in Figure 7.1). The majority of these circuits are examined in detail in many volumes of electronic engineering, reason for us to make only short presentations here and provide references to deepen knowledge about these items.

## 9.1 Signal Amplifiers

The signal amplification and conditioning in ISs are currently almost exclusively made using integrated amplification devices. The exceptions are very rare and have no other justification such as low-cost.

First, amplifiers can be classified based on the controlled sources they are designed to realize.

The four main categories correspond to the four types of ideal voltage/current controlled voltage/current sources. Correspondingly, they have infinite or zero input impedance and zero or infinite output impedance. Hence, source and load impedances have no effect on the input-output relationships of these ideal circuits.

Practical amplifiers, however, have finite or nonzero input/output impedances, respectively. Thus, the size of amplifier impedance levels relative to load and source impedances has to be considered.

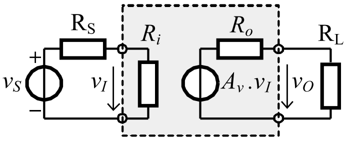
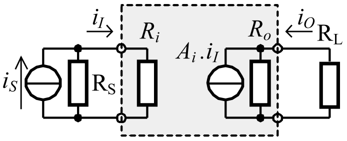
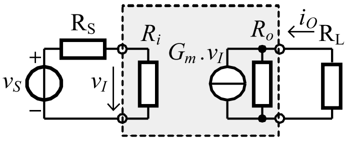
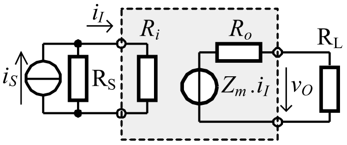
For a brief presentation, the characteristics of ideal and practical amps are summarized in Table 9.1.

The transfer functions in Table 9.1 are

- Voltage amplification $A_v$;
- Current amplification $A_i$;
- Transconductance $G_m$;
- Transimpedance $Z_m$.

These functions are nondimensional (voltage and current amplifications) or have a conductance or impedance character.

**Table 9.1** Characteristics of Four Main Categories of Amplifiers (Ideal and Real)

| *Amplifier Type* | | *Input/Output Impedence* | *Transfer Ratio (Gain)* | *Circuit Model* |
|---|---|---|---|---|
| Voltage amplifier | Ideal | $R_i \to \infty$<br>$R_o \simeq 0$ | $v_o = A_v \cdot v_s$ |  |
| | Real | $R_i >> R_S$<br>$R_o << R_L$ | $v_o \simeq A_v \cdot v_s$ | |
| Current amplifier | Ideal | $R_i \simeq 0$<br>$R_o \to \infty$ | $i_o = A_i \cdot i_s$ |  |
| | Real | $R_i << R_S$<br>$R_o >> R_L$ | $i_o \simeq A_i \cdot i_s$ | |
| Transconductance amplifier | Ideal | $R_i \to \infty$<br>$R_o \to \infty$ | $i_o = G_m \cdot v_s$ |  |
| | Real | $R_i >> R_S$<br>$R_o >> R_L$ | $i_o \simeq G_m \cdot v_s$ | |
| Transimpedance amplifier | Ideal | $R_i \simeq 0$<br>$R_o \simeq 0$ | $v_o = Z_m \cdot i_s$ |  |
| | Real | $R_i << R_S$<br>$R_o << R_L$ | $v_o \simeq Z_m \cdot i_s$ | |

## 9.1.1 Operational Amplifiers: Definition and Applications

Today, the most popular amplifier is still the operational amplifier (op-amp), which is a direct-coupled, high-gain *voltage amplifier* (voltage-controlled voltage source) used to perform a wide variety of functions and applications. The name op-amp was established in the 1960s even though the device is much older. Jung [1] describes the three phases of the op-amp history: vacuum tube op-amps, solid-state hybrid op-amps, and IC op-amps.

Many op-amps comprise a cascade of differential pairs (see Figure 8.24) followed by a gain stage, a voltage-level shifter, and finally, by a complementary emitter-follower output stage [2, 3]. The schematic symbol of an op-amp and its equivalent circuit are shown in Figure 9.1.

The + and – symbols at the op-amp input refer to the noninverting and inverting input terminals. The output voltage $v_{\mathrm{OUT}}$ is the amplified difference signal:

$$v_{\mathrm{OUT}} = A_v \cdot v_{\mathrm{IN}} = A_v \cdot \left(v_{\mathrm{IN+}} - v_{\mathrm{IN-}}\right) \tag{9.1}$$

If $v_{\mathrm{IN+}} = 0$, $v_{\mathrm{OUT}}$ is 180° out of phase (inverted) with respect to the input signal $v_{\mathrm{IN-}}$. In contrast, if $v_{\mathrm{IN-}} = 0$, the output $v_{\mathrm{OUT}}$ and the input $v_{\mathrm{IN+}}$ are in phase (not inverted).

The evaluation of the application circuits with op-amps is generally performed using the concept ideal op-amp, characterized by the following 5-attributes model:

1. The input resistance: $R_i \to \infty$ (open-circuited). Consequently, no current enters either input terminal.
2. The output resistance: $R_o = 0$.
3. The voltage gain $A_v \to \infty$. As $A_v$ is infinite but the output voltage is finite, it requires $v_{\mathrm{IN}} = 0$.
4. The amplifier responds equally to all frequencies (the bandwidth is infinite).
5. When $v_{\mathrm{IN-}} - v_{\mathrm{IN+}} = 0$, $v_{\mathrm{OUT}} = 0$. The converse is also true.

In contrast to ideal op-amps, practical commercial op-amps have a high but definite input resistance and a nonzero low value output resistance. A significantly higher input resistance is offered by the bipolar op-amps with JFET input transistors from the series TI08xx (manufacturer Texas Instruments Inc.). The voltage gain is finite and has a specified frequency sweep. When $v_{\mathrm{IN-}} = v_{\mathrm{IN+}}$, $v_{\mathrm{OUT}}$ is different from zero (input offset voltage).

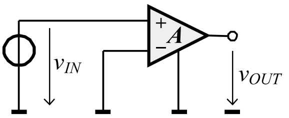
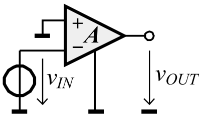
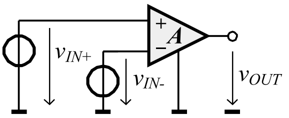
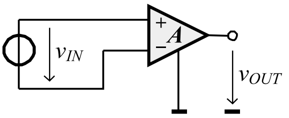
The base amplifier configurations with op-amps are represented in Table 9.2.

Implementing a negative feedback, these typical configurations can be improved by reason of improved stability, reduced magnitude of spurious signals, increased bandwidth, and operating speed, diminished nonlinearities.

The elementary feedback topologies and their main characteristics are summarized in Table 9.3.

Beside these applications, the op-amps are often implemented in inverting or noninverting summing amplifiers (adders), analog adder-subtractor amplifiers, integrators and differentiators, current-to-voltage or voltage-to-current convertors, as well as in nonlinear circuits (e.g., exponential and logarithmic amplifiers, which use the nonlinear v-to-i characteristic of bipolar transistors; see Section 9.4) [4].

**Table 9.2** Amplifiers with Ideal Operational Amplifiers

| *Amplifier Type* | *Schematic* | *Transfer Function* |
|---|---|---|
| Noninverting amplifier |  | $v_{\mathrm{OUT}} = A \cdot v_{\mathrm{IN}}$ |
| Inverting amplifier |  | $v_{\mathrm{OUT}} = -A \cdot v_{\mathrm{IN}}$ |
| Differential amplifier |  | $v_{\mathrm{OUT}} = A \cdot (v_{\mathrm{IN+}} - v_{\mathrm{IN-}})$ |
| Differential amplifier<br>(floating inputs) |  | $v_{\mathrm{OUT}} = A \cdot v_{\mathrm{IN}}$ |

The adder-subtractor with equalized amplifications (Figure 9.2(a)) provides the following transfer function:

$$v_{\mathrm{OUT}} = \frac{R_R}{R_I} \cdot \left(v_{\mathrm{IN1}} + v_{\mathrm{IN2}} - v_{\mathrm{IN3}} - v_{\mathrm{IN4}}\right) \tag{9.2}$$

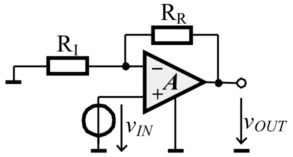
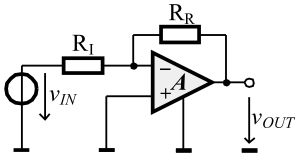
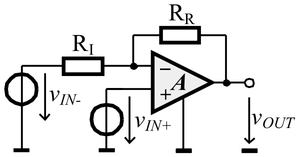
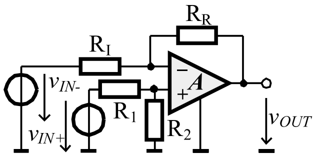
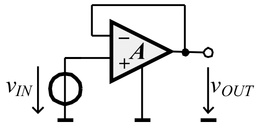
The output of the integrator (Figure 9.2(b)) is expressed by:

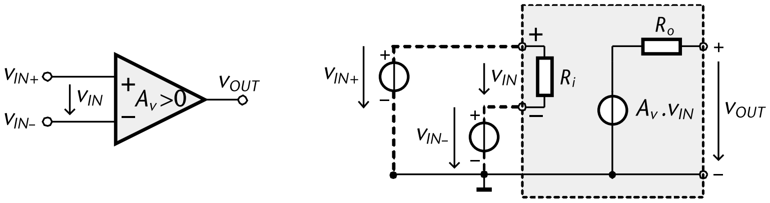
$$v_{\mathrm{OUT\_}I} = -\frac{1}{RC}\int v_{\mathrm{IN}}(t)\,dt \tag{9.3}$$

whereas the output of the differentiator (Figure 9.2(c)) is expressed by:

$$v_{\mathrm{OUT\_}D} = -RC\frac{dv_{\mathrm{IN}}(t)}{dt} \tag{9.4}$$

**Table 9.3** Negative-Feedback Amplifiers with Operational Amplifiers

| *Amplifier Type* | *Schematic* | *Reverse Transmission β of the Feedback Network (8.44) Gain with Feedback (True and Round if A >>1) (see 8.45)* |
|---|---|---|
| Noninverting amplifier |  | $\beta_N \overset{\mathrm{def}}{=} \frac{v_{\mathrm{OUT}}}{v_-} = \frac{R_I}{R_I + R_R}$<br>$\frac{v_{\mathrm{OUT}}}{v_{\mathrm{IN}}} = \frac{A}{1 + A\beta_N} \approx 1 + \frac{R_R}{R_I}$ |
| Inverting amplifier |  | $\beta_I \overset{\mathrm{def}}{=} \frac{v_{\mathrm{OUT}}}{v_-} = \frac{R_I}{R_R}$<br>$\frac{v_{\mathrm{OUT}}}{v_{\mathrm{IN}}} = \frac{-A}{1 + (A+1)\beta_I} \approx -\frac{R_R}{R_I}$ |
| Differential amplifier |  | $v_{\mathrm{OUT}} \approx \frac{1}{\beta_N}v_{\mathrm{IN+}} - \frac{1}{\beta_I}v_{\mathrm{IN-}}$<br>(for $\beta_N$ and $\beta_I$ see expressions above in this column) |
| Differential amplifier with equalized amplifications (subtractor) |  | when: $\frac{R_R}{R_I} = \frac{R_2}{R_1}$<br>$v_{\mathrm{OUT}} \approx \frac{R_R}{R_I}\left(v_{\mathrm{IN+}} - v_{\mathrm{IN-}}\right)$ |
| Voltage follower (impedance convertor) |  | $\beta \overset{\mathrm{def}}{=} \frac{v_{\mathrm{OUT}}}{v_-} = 1$<br>$\frac{v_{\mathrm{OUT}}}{v_{\mathrm{IN}}} = A_F = \frac{A}{1 + A\beta} \approx 1$ |

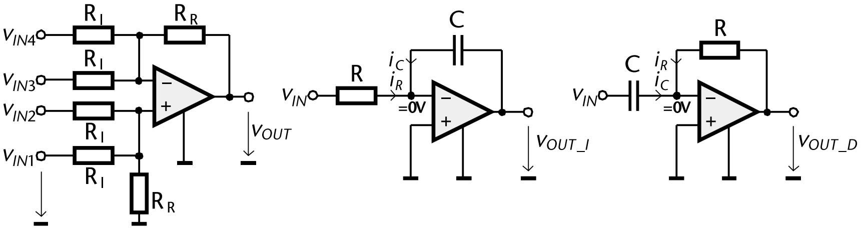

$v_{IN+}$
$v_{IN-}$
$v_{IN}$
$A_v>0$
$v_{OUT}$
$R_i$
$R_o$
$A_v.v_{IN}$
+
−

(a) (b)

**Figure 9.1** The operational amplifier: (a) circuit symbol, and (b) equivalent circuit.

An important op-amps application for sensor engineering (capacitive, reactance variation, etc.) is the instrumentation amplifier, which provides an accurate processing of the output signal of measuring bridges (DC or AC). The circuit consists of three op-amps. It realizes a high impedance signal extraction of a measuring bridge and converts the differential signal into an asymmetric, GND-related output signal (Figure 9.3). The circuit has an additional input, which optionally can be supplied with a voltage $v_{\text{REF}}$ for an offset compensation or is grounded.

The input stage is fully differential (differential input and output). The gain of buffers $A_1$ and $A_2$ is unity for common-mode voltage but is high for a difference signal.


$v_{IN4}$
$v_{IN3}$
$v_{IN2}$
$v_{IN1}$
$R_I$
$R_R$
$v_{OUT}$
R
C
$i_C$
$i_R$
=0V
$v_{IN}$
$v_{OUT\_I}$
$v_{OUT\_D}$

(a) (b) (c)

**Figure 9.2** Schematic diagrams of special op-amp applications: (a) adder-subtractor with four equalized inputs, (b) integrator, and (c) differentiator.

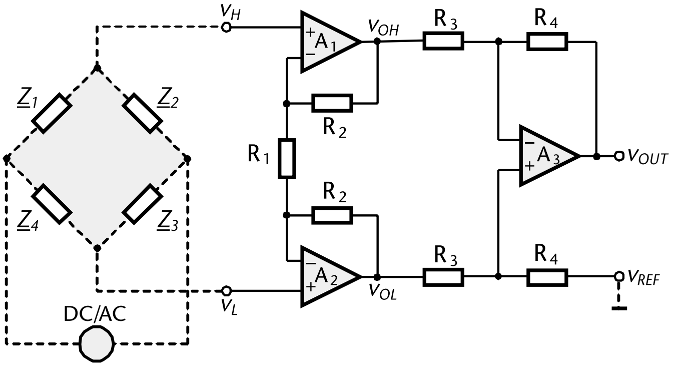

$v_H$
$v_{OH}$
$A_1$
$A_2$
$A_3$
$R_1$
$R_2$
$R_3$
$R_4$
$v_{OUT}$
$v_{REF}$
$v_{OL}$
$v_L$
$\underline{Z}_1$
$\underline{Z}_2$
$\underline{Z}_3$
$\underline{Z}_4$
DC/AC

**Figure 9.3** High performance instrumentation amplifier (version with three op-amps) evaluating the output signal of sensor measuring bridge ($R_3$ and $R_4$ must be matched).

Indeed, due to the third attribute of the five-attributes model of the ideal op-amp mentioned above, the voltage drop across the resistor $R_1$ is equal to the difference $v_H - v_L$. If a common-mode signal is under consideration, $v_H = v_L$ and the voltage across $R_1$ is zero. Hence, there is no current in $R_1$ and $R_2$. Consequently, $v_{OH} = v_H$ and $v_{OL} = v_L$ and the buffers act as unity gain amplifiers.

However, if $v_H - v_L$ is different from zero, there is current in $R_1$ and $R_2$ and $(v_{OH} - v_{OL}) > (v_H - v_L)$, which is an existing differential gain.

The output stage ($A_3$) is a difference amplifier converting a differential voltage into a single-ended output voltage [1, 5, 6]. The gain equation for the entire circuit with grounded reference voltage input is:

$$A = \frac{v_{\text{OUT}}}{v_H - v_L} = \left(1 + 2\frac{R_2}{R_1}\right) \cdot \frac{R_4}{R_3} \tag{9.5}$$

and can be conveniently adjusted by the single-element $R_1$.

The schematic has a symmetrical structure, which is easy to be designed and tested. It is a key piece to evaluate the voltage in the diagonal branch of measuring DC or AC bridges (see the common bridge $\underline{Z}_1$ to $\underline{Z}_4$, dashed lines represented in Figure 9.3 and supplied at opposite vertices by an optional DC or AC source). Such bridges with four terminals are described in previous sections (e.g., Figures 3.10 and 3.16).

The instrumentation amplifier can be realized with three op-amps ($A_1$ and $A_2$ should be matched) or implemented in specific integrated circuits. For example, the input stage in Figure 9.3 is available in IC form in LT1101AC (Linear Technology), INA114A (Texas Instruments). The classical circuits AD621-AD624 (Analog Devices) offer flexible and performant fully integrated instrumentation amplifiers.

Monolithic (single-chip) amplifiers are designed to have very high differential input resistances (>100 MΩ) to minimize loading effects on the measurement system. In addition, they have extremely high common-mode rejection ratios (CMRR) (in the order of 120 dB) because they have to amplify very small difference signals ($\approx$10 $\mu$V) in the presence of relatively high common-mode signals (in the order 1V).

### 9.1.2 Operational Amplifiers: Frequency Response, Stability, and Compensation

#### 9.1.2.1 Frequency Response, Bode Plot

The frequency of the signal provided by the inductive sensing elements, which has to be amplified, displays wide variation (from DC up to the range of a few megahetz). Hence, to evaluate the result of the amplification, it is necessary to know the transfer function of the amplifier $A(s)$ at each frequency.

A convenient method by which this information is obtained is the frequency-response characteristic, which is the plots of the magnitude $|A(j\omega)|$ and the phase of the $A(j\omega)$ versus $\omega$. Usually, $|A(j\omega)|$ is expressed in decibels (dB) and given by:

$$|A(j\omega)| \text{ in dB} = 20\log|A(j\omega)| \tag{9.6}$$

When $|A(j\omega)|$ in decibels is plotted, the frequency-response characteristic is called the Bode diagram. Determining the Bode diagram by algebraic manipulation is a moderate chore. For many purposes, it is adequate to draw an approximate characteristic, called the asymptotic Bode diagram [4].

Starting with the general complex expression, an amplification:

$$A(s) = A_0 \frac{1 + a_1 s + a_2 s^2 + \ldots + a_m s^m}{1 + b_1 s + b_2 s^2 + \ldots + b_n s^n} \tag{9.7}$$

substituting $s$ by $j\omega$ and factoring the numerator and denominator polynomials results in:

$$A(j\omega) = A_0 \frac{\prod_1^m \left(1 + \frac{s}{z_i}\right)}{\prod_1^n \left(1 + \frac{s}{p_j}\right)} = A_0 \frac{\prod_1^m \left(1 + \frac{j\omega}{z_i}\right)}{\prod_1^n \left(1 + \frac{j\omega}{p_j}\right)} \tag{9.8}$$

where the roots $-z_i$ and $-\mathrm{p}_j$ of the numerator and denominator, respectively, are called zeros and poles (see 8.51).

The magnitude-curve portion of the Bode diagram plotted in decibels is a logarithmic function of the expression (9.8) and thus the products become sums:

$$A(j\omega) \text{ in dB} \equiv 20\log A_0 + \sum_1^m 20\log\left(1 + \frac{j\omega}{\omega_{0i}}\right) - \sum_1^n 20\log\left(1 + \frac{j\omega}{\omega_{0j}}\right) \tag{9.9}$$

In conclusion, the magnitude-curve portion of the Bode diagram is the result of a graphical addition of the terms from (9.9). These parts can be approximately represented as follows:

- At low frequencies ($\omega/\omega_{0i} << 1$ and $\omega/\omega_{0j} << 1$), the magnitude of partial terms is:

$$\left|1 + \frac{j\omega}{\omega_{0i}}\right| \approx \left|1 + \frac{j\omega}{\omega_{0j}}\right| \approx 1 \tag{9.10}$$

and therefore:

$$20\log\left|1 + \frac{j\omega}{\omega_{0i}}\right| \approx 20\log\left|1 + \frac{j\omega}{\omega_{0j}}\right| \approx 0\ [\mathrm{dB}] \tag{9.11}$$

- At high frequencies ($\omega/\omega_{0i} >> 1$ and $\omega/\omega_{0j} >> 1$), the functions become:

$$1 + \frac{j\omega}{\omega_{0i}} \approx \frac{j\omega}{\omega_{0i}} \text{ and } \frac{1}{1 + \frac{j\omega}{\omega_{0j}}} \approx -\frac{j\omega_{0j}}{\omega} \tag{9.12}$$

and therefore:

$$20\log\left|1+\frac{j\omega}{\omega_{0i}}\right| \approx 20\log\frac{\omega}{\omega_{0i}} \ [\text{dB}] \tag{9.13}$$

and

$$20\log\frac{1}{\left|1+\dfrac{j\omega}{\omega_{0j}}\right|} \approx -20\log\frac{\omega}{\omega_{0j}} \ [\text{dB}] \tag{9.14}$$

Since a factor of 10 between the actual frequency $\omega$ and a corner frequency $\omega_{0i}$ or $\omega_{0j}$ is a linear increment on the logarithmic frequency scale, (9.13) and (9.14) are straight lines on the Bode plots with slopes 20 dB/decade and −20 dB/decade. According to the expression (9.9), the asymptotic Bode plot can be sketched by graphically adding these straight lines.

*Example 9.1*

Let us consider a hypothetical op-amp is characterized by the following features: voltage amplification at low frequencies $A_{dc}$ = 10,000, the corner frequencies $f_1 = \omega_1/2\pi$ = 1 kHz and $f_2 = \omega_2/2\pi$ = 100 kHz and a complex amplification with two poles expressed by:

$$A(j2\pi f) = \frac{A_{dc}}{\left(1+j\dfrac{f}{f_1}\right)\left(1+j\dfrac{f}{f_2}\right)} \tag{9.15}$$

The asymptotic Bode plot of the op-amp (solid broken line in Figure 9.4) is the graphical sum of the following three asymptotes (thin dotted lines):

1. A horizontal straight line with the ordinate 20 log10,000 = 80 dB, corresponding to the first term, namely $A_{dc}$;
2. An asymptote passing through the corner frequencies $f = f_1$ and having a slope of −20 dB per decade, corresponding to the second term;
3. An asymptote passing through the corner frequencies $f = f_2$ with the same slope.

The resulting asymptotic plot offers the following benefits:

- Estimation of the actual Bode plot. This plot is much closer to the asymptotic representation and has theoretically the larger deviation of −3 dB at the corner frequency.
- Piecewise linear approximation of the frequency response of any feedback amplifier realized with this op-amp.

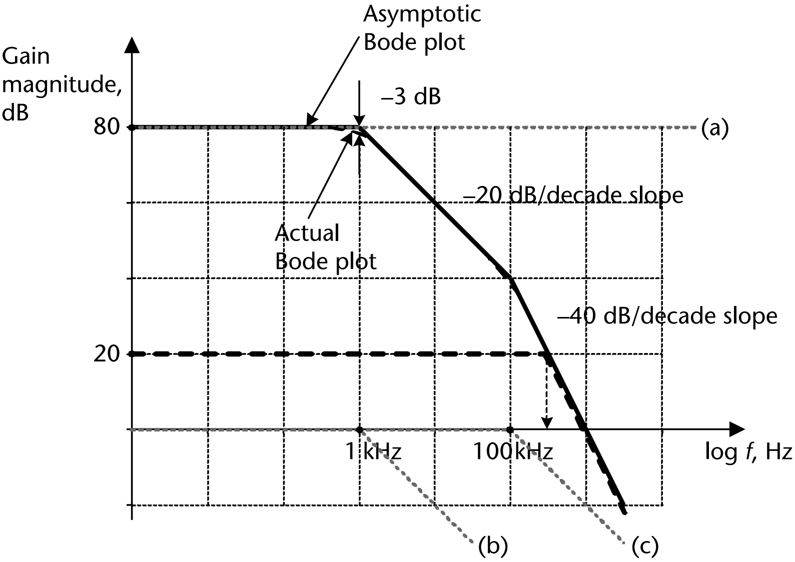


**Figure 9.4** Magnitude-curve portion of the Bode plot to Example 9.1: (a) magnitude plot of $A_{dc}$, (b) magnitude plot of 1/(1 + jf/1 kHz) and (c) magnitude plot of 1/(1 + jf/100 kHz).

Supposing the op-amp is used as a noninverting amplifier (see Table 9.3) with a frequency-independent feedback network with $R_R = 9$ kΩ and $R_I = 1$ kΩ, the low-frequency amplification becomes:

$$A_{dc,N} = \frac{A_{dc}}{1 + A_{dc}\beta_N} \approx \frac{1}{\beta_N} = 1 + \frac{R_R}{R_I} = 10 \tag{9.16}$$

This computation is sufficient to represent the frequency response of the feedback amplifier. The diagram starts at low frequency with the horizontal dashed line having the ordinate 20 log $A_{dc,N}$ = 20 log 10 = 20 dB.

The dashed line is pulled through until the intersection with the op-amp Bode plot. This intersection defines a frequency called bandwidth $B$. After this resulting corner frequency of the feedback amplifier (320 kHz), its Bode plot follows the op-amp plot and provides:

- Immediate determination of the gain-bandwidth product;
- Graphical evaluation of the stability and actual phase margin.

### 9.1.2.2 Stability, Nyquist Criterion, and Diagram

The resulting frequency response of an op-amp, however, can be altered by the application of feedback. In some cases, the application of feedback results in an unstable system, which is a system that provides an output signal without an input signal.

According to a physical view, a system is stable if a transient disturbance of finite duration results in a response which dies out, and that should happen at all frequencies, and not merely over the frequency range of interest. This behavior is opposite to the oscillator behavior described in Section 8.2. They are deliberately designed as unstable systems and a transient disturbance produces an output, which persists indefinitely or increases until it is limited by some nonlinearities in the circuit.

The mathematical definition of the stability states that a system is stable if and only if all bounded input signals produce bounded output signals. To give an example of bounded signal, the sin-wave is bounded by unity.

In view of this definition, stability involves the study of the system transfer function. If all system poles (denominator roots) lie in the open left half (excluding imaginary axis) of the complex plane, the system is stable. Consequently, the system is stable if all poles have negative real parts.

Whether a feedback amplifier, characterized by a resulting amplification of:

$$A_F = \frac{A}{1 - \beta A} \tag{9.17}$$

for a given op-amp gain $A$ and reverse transmission of the feedback network $\beta$, is stable, is determined by the roots of the denominator. The feedback amplifier is stable when the roots of $[1 - \beta(s) \cdot A(s)]$ all lie in the open left half of the complex plane.

This general criterion of stability can be shown to be equivalent to the Nyquist criterion (1931), which is based on the Nyquist diagram (Figure 9.5)

This diagram is the line of loop gain $T = \beta A$ plotted in the complex plane for frequencies $f$ of $-\infty \leq f \leq +\infty$. The plotting for negative frequencies (dashed curve part) can be obtained by substituting $-j$ for $j$; that is, by up-down mirroring the line of $\beta A$ for positive frequencies (solid curve part) with respect to the real axis of the complex plane. For a physically realizable system, $A$ is zero at $f = -\infty$ and $f = +\infty$; hence, the Nyquist diagram is a closed curve.

Applying this diagram, the Nyquist criterion of stability can be expressed as follows: If $A$ and $\beta$ of (9.17) describe stable systems, then the system described by $A_F$ is stable if and only if the Nyquist diagram does not encircle the point (−1 +j0) on the negative real axis.

There are several descriptions of the term "encircle." For a mathematical evaluation, a vector $\underline{\mathrm{V}}$ can be drawn between the point (−1 +j0) and the moving point on the Nyquist curve. If the total angle traversed by this vector is zero as it moves along the entire curve from $f = -\infty$ and $f = +\infty$, then the loop does not encircle the point (−1 +j0) and the system is stable.

According to an intuitive explanation, the closed curve can be thought of as a loop of string with a stake driven into the complex plane at the (−1 +j0) point. If the loop of string can be removed without lifting it over the stake, then it does not encircle the stake and the system is stable.

For example, the Nyquist loop in Figure 9.5(a) does not encircle the point (−1 +j0) and the system is stable. In contrast, the loop in Figure 9.5(b) does encircle the point (−1 +j0), which denotes an unstable system.

The Nyquist criterion is a useful tool for determining the limits of stability, and safety margins should be provided by staying away from these limits.

Figure 9.6 shows a Nyquist diagram part for positive frequencies and a unit circle, corresponding to $T(j\omega) = \beta(j\omega) \cdot A(j\omega) = 1$ (0 dB).

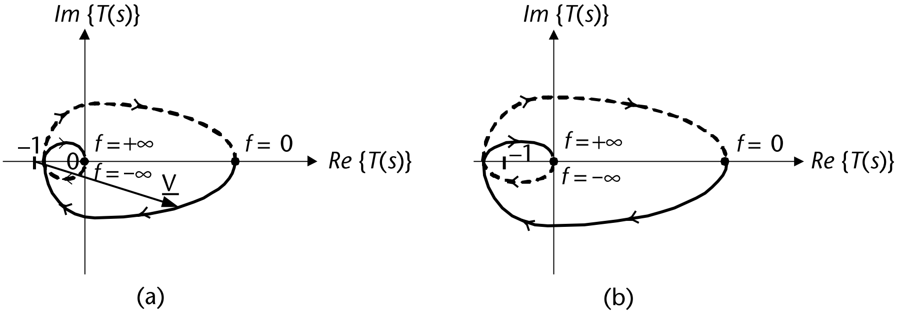


**Figure 9.5** Nyquist diagrams of (a) a stable amplifier and (b) of an unstable system.

The frequency at which these graphs intersect is called gain-crossover frequency $f_g$. If the intersection point lies in the third quadrant of the complex plane, the system is stable (Figure 9.6).

An alternative approach requires that, at the frequency $f_g$ where the magnitude of the loop gain $\beta A$ becomes equal to one, the phase of this quantity should fulfill the formula:

$$\varphi_{|\beta A|=1} \geq -180^\circ + \varphi_M \tag{9.18}$$

where $\varphi_M$ is a positive phase margin that is typically between 30° and 60°.

A second significant intersection point is between the Nyquist curve and the negative real axis, corresponding to a phase of $\beta A$ equal to −180°. It defines the phase-crossover frequency $f_\varphi$ and the magnitude of $\beta A$ at this frequency represents the gain margin.

This information contained in the Nyquist diagram is often more conveniently displayed in a Bode diagram [4].

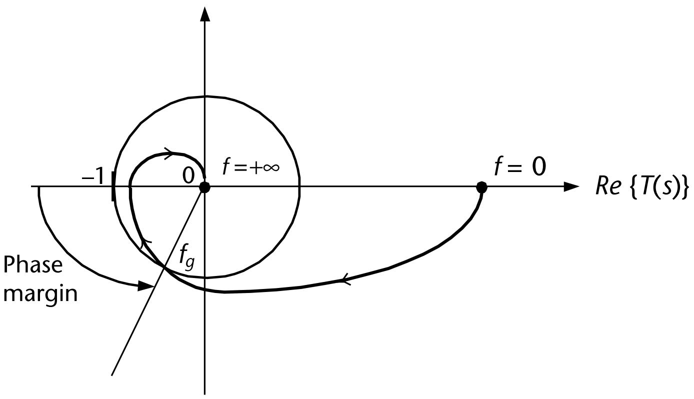


**Figure 9.6** Evaluation of the phase margin $\varphi_M$ by means of the Nyquist diagram.

*Example 9.2*

Let us continue Example 9.1 with a comparable op-amp ($A_{dc} = 1{,}000$ and identical corner frequencies) and frequency independent feedback network. The target is to graphically determine the phase margin of a resulting noninverting amplifier with $A_{F,dc} = 10$, by means of asymptotic Bode plots.

The loop gain:

$$T(j2\pi f) = \beta(j2\pi f)\cdot A(j2\pi f) = \frac{1{,}000}{(1 + jf/1\text{ kHz})(1 + jf/100\text{ kHz})} \tag{9.19}$$

has its Bode plots displayed in Figure 9.7.

The results are immediately provided by the arrow in the plots. The amplifier is stable because it has a positive phase margin of $\varphi_M \approx 60°$.

### 9.1.2.3 Compensation Techniques

The stability of a feedback amplifier can be achieved by implementing compensation techniques of the frequency response provided by hardware. In other words, the compensation consists of the inclusion of a frequency dependent network in the loop $\beta A$, either in the op-amp or in the feedback network. The first option is much more widespread and it refers to the internally compensated op-amps with a dominant pole at low frequencies of 10 to 100 Hz (e.g., the LM13741, manufacturer National Semiconductor/Texas Instruments Inc.) [4]. These op-amps offer the unconditional stability of their feedback applications but have a drastically reduced open-loop bandwidth. In contrast, the uncompensated op-amps need an external compensation measure; however, they are more flexible in terms of design and resulting performances.

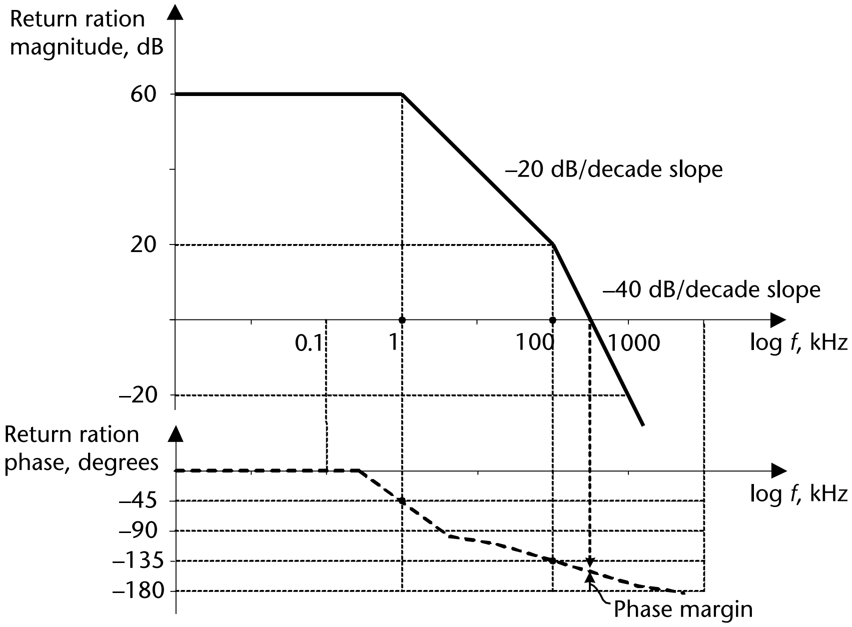


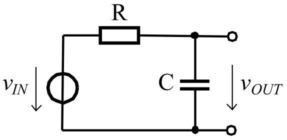
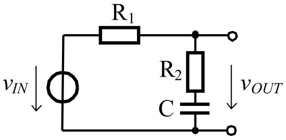
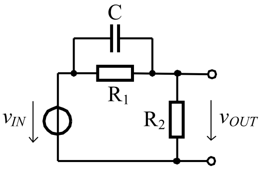
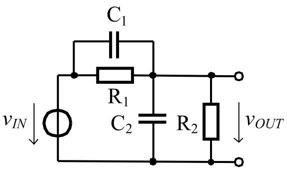
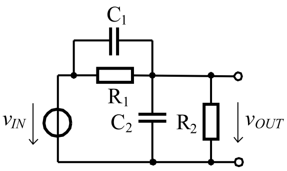
**Figure 9.7** Magnitude curve and phase curve of the Bode diagram to Example 9.2; phase margin.

The well-established compensation techniques are summarized in Table 9.4. The listed transfer functions evidence whether the output legs or leads the input. The compensation networks can be also combined. For example, Figure 9.8 shows the implementation of the lead-leg compensation into a noninverting and an inverting feedback amplifier.

The feedback return $\beta_N$ in Figure 9.8(a) is:

$$\beta_N = \frac{R_I}{R_F + R_I} \cdot \frac{1 + \dfrac{jf}{f_L}}{1 + \dfrac{jf}{f_U}} \tag{9.20}$$

**Table 9.4** Schematics and Transfer Function of Classical Compensation Networks

| *Compensation Type* | *Schematic* | *Transfer Function Expression*<br>*Resulting Corner Frequencies* |
|---|---|---|
| Lag compensation |  | $\frac{v_{\mathrm{OUT}}(f)}{v_{\mathrm{IN}}(f)} = \frac{1}{1 + jf/f_0}$;<br>$f_0 = \frac{1}{2\pi RC}$ |
| Modified lag compensation |  | $\frac{v_{\mathrm{OUT}}(f)}{v_{\mathrm{IN}}(f)} = \frac{1 + jf/f_2}{1 + jf/f_1}$;<br>$f_1 = \frac{1}{2\pi(R_1 + R_2)C}$<br>$f_2 = \frac{1}{2\pi R_2 C}$ |
| Lead compensation |  | $\frac{v_{\mathrm{OUT}}(f)}{v_{\mathrm{IN}}(f)} = \frac{R_2}{R_1 + R_2}\frac{1 + jf/f_1}{1 + jf/f_2}$<br>$f_1 = \frac{1}{2\pi R_1 C}$<br>$f_2 = \frac{1}{2\pi C R_1 R_2/(R_1 + R_2)}$ |
| Voltage divider |  | $\frac{v_{\mathrm{OUT}}(f)}{v_{\mathrm{IN}}(f)} = \frac{R_2}{R_1 + R_2}\frac{1 + jf/f_1}{1 + jf/f_2}$<br>$f_1 = \frac{1}{2\pi R_1 C_1}$<br>$f_2 = \frac{1}{2\pi(C_1 + C_2)R_1 R_2/(R_1 + R_2)}$ |
| Compensated voltage divider |  | $\frac{v_{\mathrm{OUT}}(f)}{v_{\mathrm{IN}}(f)} = \frac{R_2}{R_1 + R_2}$<br>$R_1 \cdot C_1 = R_2 \cdot C_2$ |

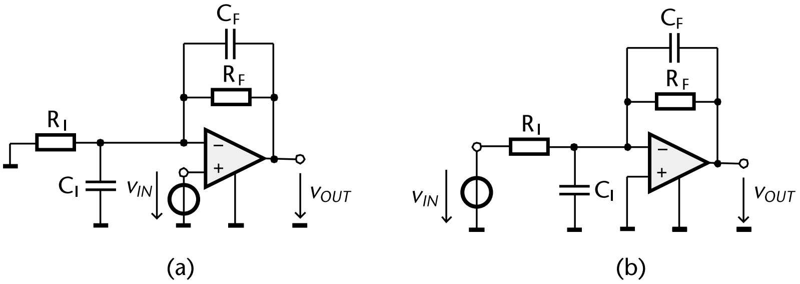


**Figure 9.8** Lead-leg compensation into (a) a noninverting and (b) an inverting feedback amplifier.

where the corner frequency $f_L$ (zero frequency) is defined by:

$$f_L = \frac{1}{2\pi R_F C_F} \tag{9.21}$$

and corner frequency $f_U$ (pole frequency) by:

$$f_U = \frac{1}{2\pi \frac{R_F R_I}{R_F + R_I}\left(C_F + C_I\right)} \tag{9.22}$$

For the lead-lag compensation, $f_L$ and $f_U$ are chosen so that $f_L < f_U$, hence the subscriptions $L$ and $U$ (lower and upper). When the DC loop gain $\beta A$ is much larger than 1, the feedback amplification can be approximated by:

$$A_{FN} \approx \frac{1}{\beta_{N,dc}} = \frac{R_F + R_I}{R_I} \tag{9.23}$$

therefore, the ratio of these corner frequencies becomes:

$$\frac{f_U}{f_L} \le A_{FN} \approx \frac{R_F + R_I}{R_I} \tag{9.24}$$

Thus, (9.24) limits the ratio of $f_U/f_L$ to the DC amplification of the feedback amplifier if the ratio $C_I/C_F$ is small.

The circuits in Figure 9.8 differ in the location of the input voltage source $v_S$. Therefore, two circuits become identical when $v_S = 0$. Since stability and phase margin considerations were independent of magnitudes of input voltages, it seems that they would be equally valid for both applications. However, a practical difference arises from the effects of nonzero impedances of the voltage sources $v_S$.

Moreover, the implementation in Figure 9.8 is the more flexible method. It can be particularized to lead or leg compensation by removing the needless parts.

## 9.2 Precision AC/DC Signal Converters

Inherent to their functional principle, the majority of ISEs are AC-excited and deliver AC output signal(s). Such signals have to be guided—directly or after an intermediate amplification—to a rectifier circuit to make a DC unipolar voltage prior to integration or filtering (Figure 9.9).

Considering an ideal input signal delivered from the ISE:

$$v_{\mathrm{IN}}(t) = V_{inm} \sin 2\pi f_{\mathrm{IN}} \cdot t \tag{9.25}$$

the output signal of the ideal half-wave rectifier is:

$$v_{\mathrm{OUT\_}HW}(t) = \begin{cases} v_{\mathrm{IN}}(t) & \text{for } 0 \le 2\pi f_{\mathrm{IN}} t \le \pi \\ 0 & \text{for } \pi \le 2\pi f_{\mathrm{IN}} t \le 2\pi \end{cases} \tag{9.26}$$

and its Fourier series representation is expressed by:

$$v_{\mathrm{OUT\_}HW}(t) = V_{inm}\left\{\frac{1}{\pi} + \frac{1}{2}\sin\left(2\pi f_{\mathrm{IN}} t\right) - \frac{2}{\pi}\left[\frac{\cos\left(4\pi f_{\mathrm{IN}} t\right)}{2^2 - 1} + \frac{\cos\left(8\pi f_{\mathrm{IN}} t\right)}{4^2 - 1} + \dots\right]\right\} \tag{9.27}$$

Similarly, the output signal of the ideal full-wave rectifier is:

$$v_{\mathrm{OUT\_}FW}(t) = \begin{cases} v_{\mathrm{IN}}(t) & \text{for } 0 \le 2\pi f_{\mathrm{IN}} t \le \pi \\ -v_{\mathrm{IN}}(t) & \text{for } \pi \le 2\pi f_{\mathrm{IN}} t \le 2\pi \end{cases} \tag{9.28}$$

where its Fourier series representation is expressed by:

$$v_{\mathrm{OUT\_}FW}(t) = V_{inm}\left\{\frac{2}{\pi} - \frac{4}{\pi}\left[\frac{\cos\left(4\pi f_{\mathrm{IN}} t\right)}{2^2 - 1} + \frac{\cos\left(8\pi f_{\mathrm{IN}} t\right)}{4^2 - 1} + \dots\right]\right\} \tag{9.29}$$

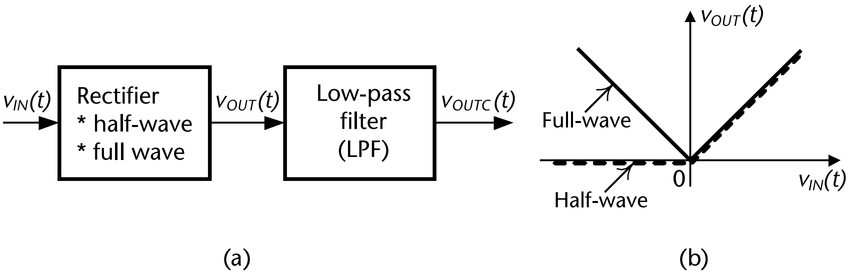


**Figure 9.9** (a) Block diagram of a precision AC/DC signal converter, and (b) ideal transfer characteristic of a half-wave and a full-wave rectifier, respectively.

If the low-pass filter has a cutoff frequency $f_C$ lower then $f_{IN}$, it passes the DC component and attenuates the harmonic components. Correspondingly, the output signal of the ideal AC/DC convertor has the expression:

$$v_{OUTC}(t) = V_{OUTC} \approx \begin{cases} \dfrac{V_{inm}}{\pi} & \text{for the half-wave rectifier} \\ 2 \cdot \dfrac{V_{inm}}{\pi} & \text{for the full-wave rectifier} \end{cases} \tag{9.30}$$

A benchmark between both possibilities immediately shows the benefits offered by the full-wave rectifier: double value of the DC output voltage $V_{OUTC}$ and better rejection of the harmonics due to the larger distance between the frequencies of the harmonics and the filter cutoff frequency.

### 9.2.1 Precision Rectifiers

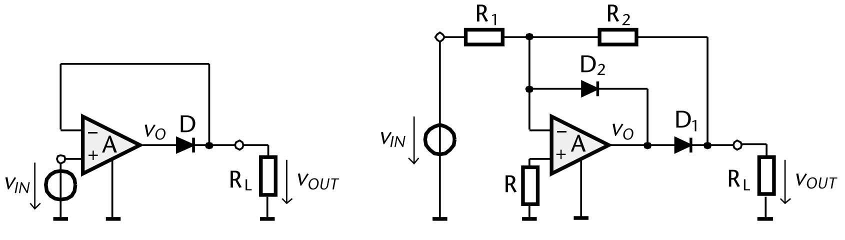
Coarse sensors use a diode or a diode bridge to rectify the ISE output signal, but diodes are not suitable for precision applications because they have a nonlinear v-to-i characteristic. Their forward voltage $V_{TH}$ (0.2 to 0.6V) is too high in comparison to many ISE output voltages and they are temperature sensitive and poorly regulated.

These problems are overcome by the use of active rectifiers, which have the diode(s) placed in their feedback circuits. By placing the diode in the feedback loop of an op-amp, the cut in diode voltage $V_\gamma$ is divided by the open-loop gain $A_V$ of the amplifier. Hence, $V_\gamma$ is virtually eliminated and the diode approaches the ideal rectifying component.

#### 9.2.1.1 Half-Wave Rectifiers

If in Figure 9.10(a) the input $v_{IN}$ goes positive by at last $V_\gamma/A_V$, then $v_O$ exceeds $V_\gamma$ and the diode D conducts. Due to the feedback with a diode in ON-state, the circuit acts a voltage follower (see Table 9.3) for positive signals (in excess of $0.6\ /\ 10^5\text{V} = 60\ \mu\text{V}$) and $v_{OUT} = v_{IN}$.

When $v_{IN}$ swings negatively, D goes OFF and no current is delivered to the external load $R_L$ (except for the small bias current of the inverting op-amp input and the diode reverse saturation current).



(a) (b)

**Figure 9.10** Active half-wave rectifier: (a) classical noninverting version, and (b) improved half-wave inverting configuration.

The circuit can be modified to rectify the negative half-wave by reversing the diode connection. The disadvantages appear when the feedback loop is open. The saturation of the op-amp shakes up the frequency response and the op-amp works with a large differential input voltage and has a high output impedance.

To prevent these effects and to speed up the rectifier, the noninverting topology is replaced by the inverting op-amp application, which contains a second diode $D_2$ (Figure 9.10(b)) [7].

If the input $v_{IN}$ goes negative, $D_1$ is ON, $D_2$ is OFF, and the circuit behaves as an inverting amplifier, so that:

$$v_{OUT}(t) \approx -\frac{R_2}{R_1} v_{IN}(t), \ \left(v_{IN} < 0\right) \tag{9.31}$$

where [7] proposes $R_1 = R_2 = 20\ \mathrm{k}\Omega$.

If $v_{IN}$ is positive, $D_1$ is OFF and $D_2$ is ON. Because of the feedback through $D_2$, a virtual ground exists at the input. The op-amp output voltage $v_O$ is negative and equal to the voltage drop on $D_2$. The rectifier output is $v_{OUT} = 0$.

The principal limitation of the circuit is the slew rate of the op-amp. As the input passes through zero, the op-amp output $v_O$ must change as quickly as possible from +0.6 to −0.6V and vice versa to switch from one diode to the other. Supposing the slew rate is 1V/$\mu$s, the switching time is 1.2 $\mu$s. Hence, 1.2 $\mu$s must be a small fraction of the period of the input sinusoid. Thus, op-amps without internal compensation (dominant pole) could provide high accuracy in a range up to of few 100 kilohertz. The input impedance is moderate.

The similar configuration but for negative output voltage can be achieved by reversing the diodes. An alternative noninverting configuration to Figure 9.10(b) is to ground the resistor $R_1$, to reverse the diodes and to apply $v_{IN}$ at the noninverting terminal.

### 9.2.1.2 Full-Wave Rectifiers

The system shown in Figure 9.11(a) gives full-wave rectification without global inversion and with a gain controllable by the one resistor $R_1$.

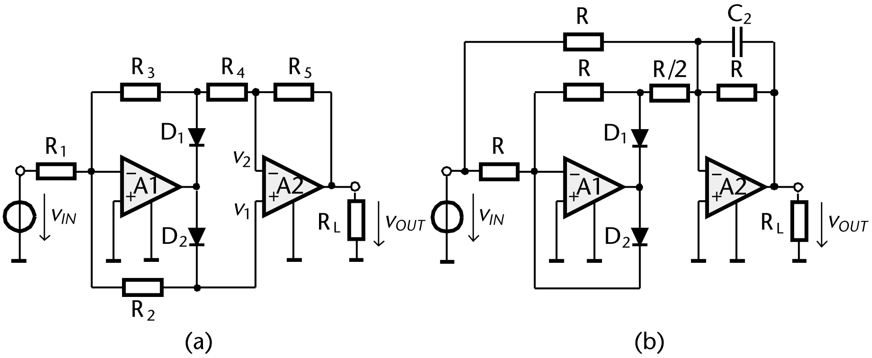


Figure 9.11 (a) Absolute value rectifier [8] and (b) Active full-wave rectifier and filter [4].

If $v_{IN}$ is positive, $D_1$ is ON and $D_2$ is OFF. Since $D_1$ conducts, the loop of $A_1$ is closed and there is a virtual ground at the input of $A_1$. Because $D_2$ is nonconducting and there is no current in the resistor $R_2$, which is connected to the noninverting terminal of $A_2$, then the voltage $v_1$ of the intermediate node is zero. Hence, the system consists of two op-amps in cascade and the positive output voltage results in:

$$v_{\mathrm{OUT}}(t) \approx \left(-\frac{R_3}{R_1}\right)\cdot\left(-\frac{R_5}{R_4}\right)\cdot v_{\mathrm{IN}}(t) = +\frac{R}{R_1}v_{\mathrm{IN}}(t) \tag{9.32}$$

which is valid for the traditional circuit setting [8]:

$$R_2 = R_3 = R_4 = R_5 = R \tag{9.33}$$

If the input $v_{\mathrm{IN}}$ goes negative, $D_1$ is OFF and $D_2$ is ON. Because of the virtual ground at the input to $A_2$, the voltages of the intermediate nodes are equal: $v_1 = v_2 = v$.

Since the inverting terminal of $A_1$ lies on virtual ground, three currents flow into this node. The KCL for this node gives:

$$i_{R1} + i_{R2} + i_{R3\_R4} = \frac{v_{\mathrm{IN}}(t)}{R_1} + \frac{v}{R} + \frac{v}{2R} = 0 \tag{9.34}$$

The solution of (9.34) expresses the voltage of the intermediate nodes:

$$v = -\frac{2}{3}\cdot\frac{R}{R_1}\cdot v_{\mathrm{IN}}(t) \tag{9.35}$$

The output voltage can be now expressed by means of the intermediate voltage:

$$v_{\mathrm{OUT}}(t) = v + i_{R5}\cdot R_5 \tag{9.36}$$

where $i_{R5}$ is equal to the current through the series network $R_4$ and $R_3$, because the inverting input of $A_2$ takes no current. Hence:

$$v_{\mathrm{OUT}}(t) = v + \frac{v}{2R}\cdot R = \frac{3}{2}\cdot v = -\frac{R}{R_1}\cdot v_{\mathrm{IN}}(t) \tag{9.37}$$

where use is made of (9.35). The sign of $v_{\mathrm{OUT}}$ is positive because $v_{in}$ is negative in this half-cycle (see the left segment of the characteristic $v_{\mathrm{OUT}}$-to-$v_{\mathrm{IN}}$ in Figure 9.9(b)).

The comparison of (9.32) and (9.37) leads to the conclusion that the circuit performs full-wave rectification (the output signal $v_{\mathrm{OUT}}$ is positive in both half-cycles and has the same amplification).

Note that for any input waveform $v_{\mathrm{OUT}}$ is proportional to the absolute value of the input $|v_{\mathrm{IN}}|$, which is why we call the schematic an absolute value circuit [8].

Reference [8] recommends this schematic as a specific application for the ultralow offset op-amp OP07. The value of each resistance is 10 kΩ, thus a unitary transfer function.

A similar configuration is described in [4] (Figure 9.11(b)). An integrating capacitor C is added to the circuit so that the output voltage is a DC voltage proportional to the average voltage value of the input signal.

#### 9.2.1.3 Full-Wave Rectifiers with Floating Load

The schematic in Figure 9.12 illustrates the possibility to include a full-wave rectifying diode bridge in the feedback loop of an op-amp [9, 10] to overcome the diode specific problems (see Section 9.2.1). The diode bridge is driven by a voltage-controlled current source. Thus, the current through the load resistance $R_L$ is independent of the diodes forward voltages $V_{TH}$, and for an ideal op-amp, it can be accurately expressed by:

$$i_{\mathrm{OUT}}(t) = \frac{|v_{\mathrm{IN}}(t)|}{R_{ADJ}} = \frac{|V_{inm}\sin\omega t|}{R_{ADJ}} \tag{9.38}$$

where $R_{ADJ}$ is the adjusting element.

This current is a full wave, and according to (9.29), it has an average value of

$$\bar{I}_{OUT} = \frac{2}{\pi \cdot R_{ADJ}} \cdot V_{inm} = \frac{2 \cdot \sqrt{2}}{\pi \cdot R_{ADJ}} \cdot V_{in} \tag{9.39}$$

The average current value $\bar{I}_{OUT}$ provides an average voltage value over the load resistor, which is proportional to the effective value of the input voltage $V_{in}$.

For voltages at the output of the op-amp in the range $-2V_{TH}$ to $+2V_{TH}$, the diodes are OFF and the feedback loop is open. The problems are identical to those of the half-wave rectifiers and can be minimized by similar methods.

### 9.2.2 Peak Detectors

The peak detector continuously follows the input voltage maintaining a buffer capacitor charged on the instantaneous value of this input voltage. The benefit of peak detectors results in a much simpler circuitry, which do not have any diodes.

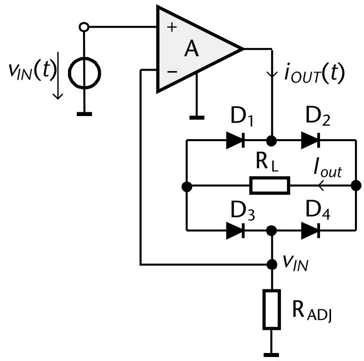


**Figure 9.12** Full-wave rectifiers with floating load.

On the other hand, the design is more complex. For confirmation, see again the circuit in Figure 7.15 and remember the two design constrictions relative to the timing network ($C_2$ in parallel to $R_4$):

- To get low ripple of the detected voltage, the capacitor must have a high capacitance and the discharging current from one peak to the next peak of the $OA_2$-input voltage has to be as low as possible;
- On the other hand, the charging and discharging times by large variations of the input has to be short enough to get high dynamic—in other words, to be able to follow the input without large time delays.

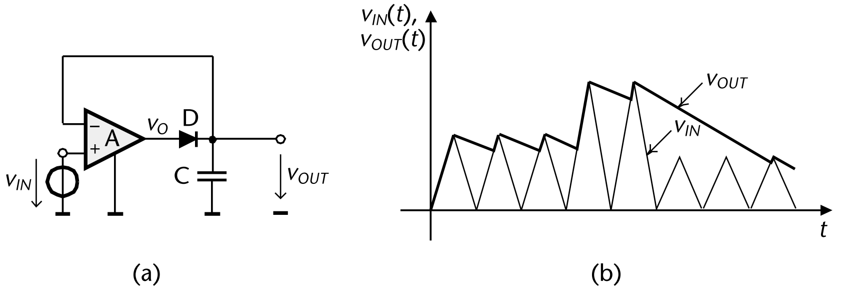
The elementary peak detector is obtained if a capacitor C is connected to the output of the active rectifier in Figure 9.10(a) and $R_L = \infty$ (Figure 9.13(a)).

If the input voltage $v_{\mathrm{IN}}$ is higher than the output voltage $v_{\mathrm{OUT}}$, the voltage at the noninverting terminal exceeds that of the inverting terminal, the op-amp output is $v_O > 0$, so that the diode D conducts. The circuit is a voltage follower and the capacitor is then charged through D (by the output current of the op-amp) to the value of the noninverting input (Figure 9.13(b)).

When $v_{\mathrm{IN}}$ falls below the capacitor voltage, the op-amp output goes negative and the diode becomes reverse-biased. The capacitor quite undecidedly discharges through the bias current of the op-amp or through the load resistance if the output is loaded. That influences the ripple of the DC output voltage but also the circuit dynamic.

To avoid this drawback and to prevent the saturation of the op-amp, a two-step improvement of the circuit in Figure 9.13(a) can be considered (Figure 9.14(a)):

- First, a low-leakage source follower with JFET is placed across the capacitor C. By connecting the inverting op-amp input to the output load, the output voltage $v_{\mathrm{OUT}}$ is forced to equal the peak value of the input voltage $v_{\mathrm{IN}}$, as desired. The capacitor voltage differs from $v_{\mathrm{OUT}}$ by the gate-to-source voltage of the JFET. For an ideal capacitor, in the hold position, its voltage changes only because of the very small JFET input current and the diode reverse current.
- A second diode $D_2$ is added to the circuit to prevent the saturation of the op-amp if $v_{\mathrm{IN}}$ falls below $v_{\mathrm{OUT}}$. In this case, $D_2$ conducts and the op-amp is a voltage follower, so that there is a virtual short circuit between terminals.



(a) (b)

**Figure 9.13** (a) An active positive peak detector. (b) An arbitrary input waveform $v_{\mathrm{IN}}$ and the corresponding output $v_{\mathrm{OUT}}$.

When $v_{IN} > v_{OUT}$, the diode $D_2$ is OFF and the circuit works like the configuration above.

To obtain a peak detector that measures the most negative value of the input voltage, it is only necessary to reverse the main diode.

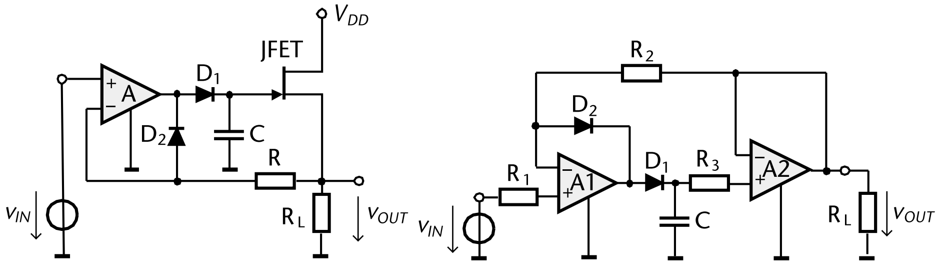
A further improvement is shown in Figure 9.14(b). The first op-amp is a peak detector with saturation avoidance (diode $D_2$) and the second op-amp works as voltage follower and should have a low input current. The resistor $R_2$ resumes the voltage difference ($v_{OUT} - v_{IN}$) when $v_{OUT} > v_{IN}$ and the resistor $R_3$ limits the surge discharging of the capacitor if the op-amp supply is switched off.

Peak detectors are economical schematics that can be realized with discrete components or standard ICs (Figures 7.3 and 7.15, respectively).

They are also suitable for integration. For example, Figure 9.15 shows a basic configuration of an integrated peak detector with a small number of parts. The integrated capacitor C (<10 pF) is charged with the current $i_{CHG}$ by the emitter follower (transistor $T_{20}$). In this example, $i_{CHG}$ is controlled with the base signal of the main transistor $T_2$ in a ring-circuit oscillator (see Figures 8.21(b) or 8.34(b)). Here, the control voltage is sine wave with a DC component expressed by (8.90).

The cyclic discharging of C between positive amplitudes occurs with constant discharging current $i_{DSC}$ through the current sink realized by the current mirror $T_{21}$ and $T_{22}$. The benefit of the constant current is a linear shape of the ripple of the detected voltage $u_C$, which makes the design flexible.

The current $i_{DSC}$ should be high enough to ensure a rapid discharging of C if the input voltage suddenly decreases but—at the same time—should be low enough to



(a) (b)

**Figure 9.14** (a) Peak detector with JFET. (b) Improved version of a peak detector with two op-amps.

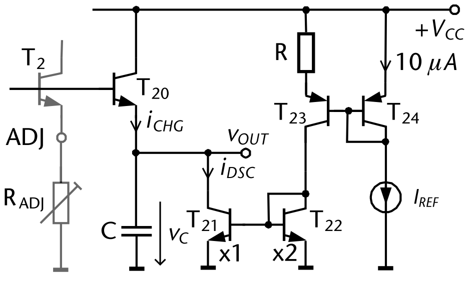


**Figure 9.15** Basic schematic of an integrated peak detector.

prevent excessive discharging of C from peak to peak, and thus a high unwanted ripple voltage. For a voltage swing $\Delta v_C$ = 1V, an integrated capacitor C = 10 pF and a working frequencies of 1 MHz, the discharging current results in a range below 1 $\mu$A.

A solution to generate such low currents is the current mirror $T_{23}$ and $T_{24}$. The reference transistor is supplied with a temperature-independent reference current $I_{REF}$ = 10 $\mu$A, which can be comfortably derived from a bandgap reference stage. The reduction of the current results in this upper current mirror by means of the adjusting resistor $R$. Applying KVL in the base loop results in:

$$2I_{DSC} \cdot R = V_{BE24} - V_{BE23} = V_T \cdot \ln\frac{10\ \mu\text{A}}{2I_{DSC}} \tag{9.40}$$

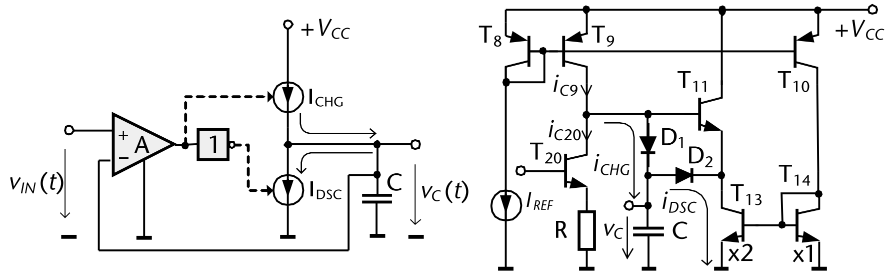
where $V_T$ is the thermal voltage (see (8.103)) and $V_{BE}$ the voltages of the matched transistors that were expressed by (8.110).

For $2I_{DSC}$ = 1 $\mu$A at 25°C, the solution of the solvable transcendent equation (9.40) [14] is $R$ = 60 k$\Omega$. This current is fed into the lower current mirror ($T_{22}$, $T_{21}$), where a halving next occurs (due to the emitter-areas ratio of 2:1). Finally, the discharge current behaves $I_{DSC}$ = 0.5 $\mu$A.

The schematic in Figure 9.15 can be improved to control both discharging and also charging current. The block diagram in Figure 9.16(a) shows an input comparator with op-amp, which compares the input signal $v_{\text{IN}}(t)$ with the voltage $v_C(t)$ across the capacitor C. The comparator toggles a DC current source and a DC current sink directly and over an inverter gate, respectively. Depending of the relationship between $v_{\text{IN}}(t)$ and $v_C(t)$, the capacitor will be charged by the enabled current source $I_{CHG}$ or discharged by the released current sink $I_{DCS}$. Obviously, the schematic realizes high-performance peak detection; it is absolutely design-flexible but more expensive.

The block diagram can be implemented in an economical configuration, which also controls the charging current (Figure 9.16(b)). This is expressed by:

$$i_{CHG} = i_{C9} - i_{C20} \tag{9.41}$$



(a) (b)

**Figure 9.16** (a) Block diagram of an integrated peak detector with controlled charging and discharging currents. (b) Economical implementation.

where $i_{C9}$ is a constant DC current delivered by the upper current mirror ($T_8$, $T_9$), which is supplied from a current reference $I_{REF}$ and $i_{C20}$ is the time-dependent collector current of $T_{20}$ (the same transistor as in Figure 9.15).

For a specific threshold amplitude $V_{lcm\text{-}th}$ of the input signal (see (8.90)), adjusted by the resistor R, the charging current becomes zero.

When $V_{lcm} < V_{lcm\text{-}th}$, $i_{DSC} > 0$ and the capacitor is charged during the positive half-wave. During the negative half-wave, $T_{11}$ goes ON and the output current of the current sink ($T_{13}$, $T_{14}$) is deviated to $V_{CC}$.

If the input achieves and exceeds the threshold $V_{lcm} > V_{lcm\text{-}th}$, the $T_{11}$ and $D_1$ go OFF and a discharging with the constant DC current:

$$i_{DSC} = 2 \cdot i_{C9} \tag{9.42}$$

is started through the $T_{13}$ and $D_2$.

The circuit has a reversed characteristic, the lowest output results for the highest input and vice versa.

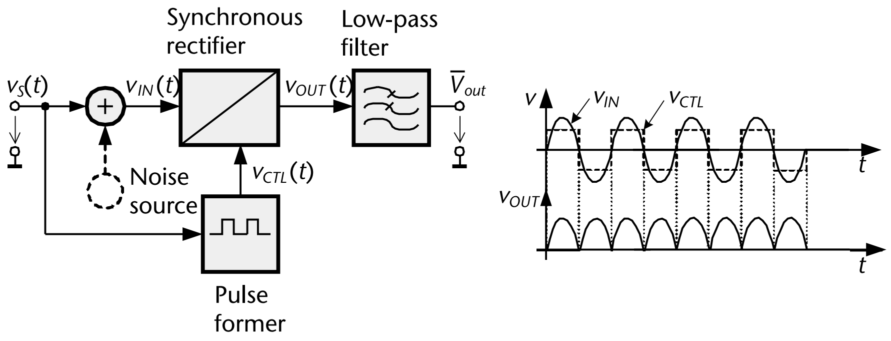
### 9.2.3 Synchronous Rectifiers

#### 9.2.3.1 Phase-Sensitive Detectors with Square-Wave Reference

An accurate AC/DC conversion can be achieved by the processing of the AC input signal $v_{IN}(t)$ in an amplifier, whose amplification has a sign, which is switched in accordance to the polarity of a control system, in particular the AC signal.

The up-to-date version in Figure 9.17.a uses a synchronous rectifier, which is an amplifier whose amplification sign can be controlled by an external chopping signal $v_{CTR}(t)$, and a low-pass filter.

Synchronous rectifiers do not use diodes or capacitors and perform a high-speed accurate AC/DC conversion. They also have specific applications as the extraction from a very noisy signal of the amplitude of those signals, which have the same frequency and phase as the control signal.



(a) (b)

Figure 9.17 (a) Block diagram of a synchronous chopping demodulator and (b) mode of operation.

The particular case for $f_{\text{IN}} = f_{CTL}$ and phase difference equal to zero is shown in Figure 9.17(b). The circuit works as a precision full-wave rectifier.

When $f_{\text{IN}} \neq f_{CTL}$ and $\varphi \neq 0$, the average voltage and the filter output $V_{\text{out}}$ progressively deviates from the theoretical value. To determine this quantity, the chopping equation is:

$$v_{\text{OUT}}(t) = v_{\text{IN}}(t) \cdot \text{CTL}(t) \tag{9.43}$$

where CTR($t$) is the control signal. Its time dependence:

$$\text{CTL}(t) = \begin{cases} +1 & \text{for } v_{CTL}(t) > 0 \\ -1 & \text{for } v_{CTL}(t) < 0 \end{cases} \tag{9.44}$$

can be expressed using the Fourier series representation of (9.44):

$$\text{CTL}(t) \approx \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{1}{2n+1} \sin(2n+1)\omega_{CTL} t \tag{9.45}$$

For an input signal having the frequency $f_{\text{IN}} = m \cdot f_{CTL}$ and for a phase difference $\varphi_m$, (9.43) becomes:

$$v_{\text{OUT}}(t) = V_{inm} \sin\left(m\omega_{CTL} t + \varphi_m\right) \cdot \frac{4}{\pi} \sum_{n=0} \frac{1}{2n+1} \sin(2n+1)\omega_{CTL} t \tag{9.46}$$

and has the solution for the average value of the output [10]:

$$\bar{V}_{out} = \begin{cases} \dfrac{2}{\pi m} V_{inm} \cdot \cos\varphi_m & \text{for } m = 2n+1 \\ 0 & \text{for } m \neq 2n+1 \end{cases} \tag{9.47}$$

with $n = 0,1,2,3....$

These expression show the fact that the circuit is selective and provides an output with the odd frequencies regarding to the control frequency. In addition, this contribution is dependent on the phase difference $\varphi_m$, and hence the synonymic term of the circuit, same-phase rectifier.

For two equal frequencies $f_{\text{IN}} = f_{CTL}$ ($m = 1$), the resulting average voltage is:

$$\bar{V}_{out} = \frac{2}{\pi} V_{inm} \text{ when } \varphi_m = 0 \tag{9.48}$$

and decreases down to zero for $\varphi_m = 90°$.

The theoretical sum of odd frequencies shown in (9.47) is valid for a low-pass filter with a time constant infinite, and consequently, with a cutoff frequency equal to zero. In practice, this is not applicable and the filter is designed for a cutoff frequency higher than $f_{\text{IN}}$.

An application of this detecting method is illustrated in Figure 9.18. The ring-circuit oscillators with LC dipole has an identical topology as in Figure 8.17. To allow the oscillator integration, the comparator is also realized with an op-amp. The op-amps are supplied with unipolar supply voltage $+V_{DD}$, and thus, the non-inverting inputs are referenced on $+V_{DD}/2$.

The oscillator operating principle is described in detail in Section 8.3.2. The equations (8.76) to (8.80) lead now to the expressions below (for symbols see these original equations).

The output signal of $\mathrm{OA}_1$ has a DC component:

$$v_{LC}(t) = \frac{V_{DD}}{2} + V_{lcm}\sin\left(\omega_0 t + \varphi_m\right) \tag{9.49}$$

This symmetrical voltage rectangular wave with defined amplitudes ($\pm V_{DD}/2$) and in phase with $v_{LC}(t)$ at the output of the $\mathrm{OA}_2$ plays a double role.

First, this voltage is converted by means the feedback adjustment resistor $R_{ADJ}$ into a symmetrical current wave:

$$i_{FB}(t) = \begin{cases} \dfrac{V_{DD}}{R_{ADJ}} & \text{for } 0 < \left(\omega_0 t + \varphi_m\right) < \pi \\ 0 & \text{for } \pi < \left(\omega_0 t + \varphi_m\right) < 2\pi \end{cases} \tag{9.50}$$

which is refed into the parallel resonant LC circuit.

The parallel resonant LC circuits for inductive sensors usually have high values of the $Q$ factor. The only effective component is the fundamental wave, with the same phase and frequency $f_{\mathrm{IN}}$ as $v_{\mathrm{LC}}(t)$ and having an amplitude according to the expression (8.79):

$$I_{fbm} = \frac{2}{\pi} \cdot \frac{V_{DD}}{R_{ADJ}} \tag{9.51}$$

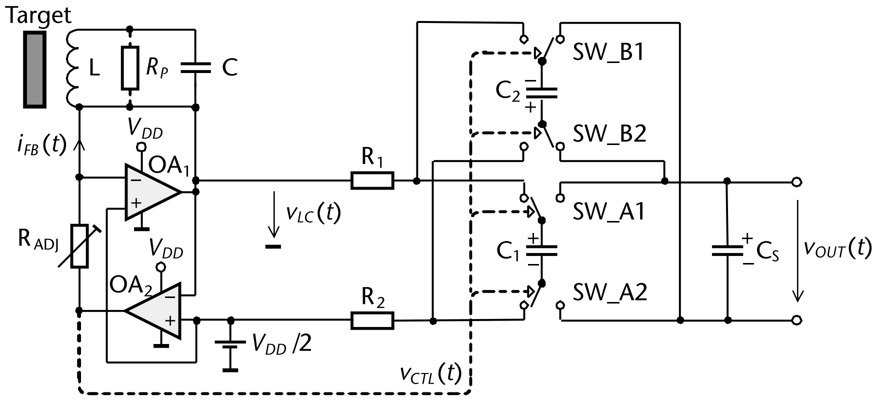


**Figure 9.18** Block schematic of an oscillator and a synchronous switching demodulator for accurate inductive proximity sensors.

This current sets the amplitude of the oscillations. From (9.49) the final oscillations expression results in:

$$v_{LC}(t) = \frac{V_{DD}}{2} + \frac{2V_{DD}}{\pi} \cdot \frac{R_P}{R_{ADJ}} \cdot \sin\left(\omega_0 t + \varphi_m\right) \tag{9.52}$$

Actually, this is the input signal of the following synchronous switching demodulator.

The second function of the OA$_2$ output voltage with the same frequency and phase as $v_{LC}(t)$ is to act as control signal for the demodulator (see signal CTL($t$) in (9.43)):

$$v_{CTL}(t) = \begin{cases} V_{DD} & \text{for } 0 < \left(\omega_0 t + \varphi_m\right) < \pi \\ 0 & \text{for } \pi < \left(\omega_0 t + \varphi_m\right) < 2\pi \end{cases} \tag{9.53}$$

The synchronous switching demodulator consists of four high-speed analog switches and three buffer capacitors. In addition, the series resistors R$_1$ and R$_2$ act together with the output capacitor C$_S$ as a low-pass filter.

If the $v_{CLT}(t) = V_{DD}$, the capacitor C$_1$ is charged on the positive amplitude peak of $v_{LC}(t)$ through the switches SW_A. The previously loaded charge in C$_2$ is transferred through the switches SW_B into the main storage capacitor C$_S$. During the next half-wave when $v_{CLT}(t) = 0$, the switches toggle, C$_2$ is charged and is C$_1$ connected to C$_S$ for charge transfer.

In this way, the output voltage $v_{\text{OUT}}(t)$ over C$_S$ is a DC voltage with a low ripple voltage. The time constant decided by R$_1$, R$_2$, and C$_S$ defines the bandwidth $B_T$ of the detector (see $B_Q$ in Table 8.1):

$$B_T = \frac{1}{\left(R_1 + R_2\right) \cdot C_S} \tag{9.54}$$

Using high-quality analog switches, the demodulator provides large operating frequency range, good linearity, and a low temperature dependence.

### 9.2.3.2 Phase-Sensitive Detectors with Analog Multiplier

The unwanted contribution of the odd frequencies (see (9.46)) can be avoided using an analog multiplier instead of a switching network [10, 11].

A rich spectrum of integrated analog multipliers is available. For example, Analog Devices offers the line of AD534, 630, 734, and 834 having part-by-part improved features. They work in one, two, or four quadrants multiplying two unipolar, unipolar/bipolar, and two bipolar voltages ($v_X$ and $v_Y$), respectively. The computation formula:

$$v_{\text{OUT}} = \frac{v_X + v_Y}{SF} \tag{9.55}$$

contains a scale factor $SF$, which can be adjusted (default value 10).

Two calculation procedures are implemented in these ICs:

1. Calculate the logarithms of the input voltages, sum the LOG values, and finally apply an ANTILOG computation of the interim sum result [13].
2. Analog multiplication by an integrated Gilbert cell multiplier [14], invented by Barrie Gilbert in the late 1960s and implemented by the majority of high-profile IC providers.

The block schematic of detection by multiplication is shown in Figure 9.19(a). The AC voltage to be rectified $v_{\mathrm{IN}}(t)$ is applied to the analog multiplier and multiplied by a sinusoidal reference voltage $v_{REF}(t) = V_{ref}\sin\omega_{REF}t$.

If $V_{ref} = SF$ (scale factor), the average value of the output results in [10]:

$$\bar{V}_{out} = \begin{cases} \frac{1}{2}V_{in}\cdot\cos\varphi_m & \text{for } f_{\mathrm{IN}} = f_{REF} \\ 0 & \text{for } f_{\mathrm{IN}} \neq f_{REF} \end{cases} \qquad (9.56)$$

In contrast to the first method, the output does not have any harmonics.

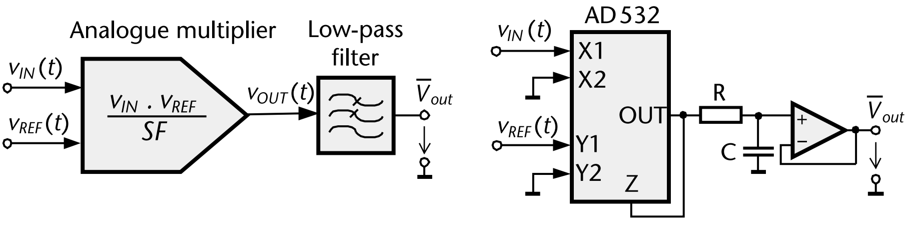
Figure 9.19(b). concretely illustrate a soft rectifier with multiplication for sinusoidal signals, having a positive output signal if the input signal is in phase with the reference signal and negative output if this is out of phase (about 180°) [15]. The average value of the output is:

$$\bar{V}_{out} = \frac{1}{20}V_{in}\cdot V_{ref}\cdot\cos\varphi_m \qquad (9.57)$$

According to [15], the output is quite independent of small phase variations (2.5° corresponds to 0.1% and 8° to 1%).

## 9.3 Sample-and-Hold Systems

Sample-and-hold (S/H) circuits realize the conversion of an analog information into a constant voltage over a certain gating-time interval. Their main application is in connection with analog-to-digital (A/D) converters. The constant S/H output may then be converted to a digital signal.



(a) (b)

**Figure 9.19** Phase-sensitive detector with analog multiplier: (a) simplified diagram, and (b) realization with the integrated circuit AD532.

The simplest form of S/H is a switch SW in series with a holding capacitor C, as shown in Figure 9.20(a) [5]. The voltage across the capacitor tracks the input signal during the time $T_{SPL}$ when a logic control gate closes SW (sampling interval), and holds the instantaneous value attained at the end of the interval $T_{SPL}$ when the control gate opens SW (for an interval $T_{HLD}$–hold interval). When the switch is closed again, the capacitor charges quickly and the output again follows the input.

The switch may be a BJT, a diode bridge, a JFET controlled by a gating-signal voltage, or a CMOS transmission gate [10]. The FET makes an excellent chopper because its offset voltage is zero.

The circuit in Figure 9.20(b) is one of the simplest practical S/H circuits. A positive pulse at the gate of N-channel FET will turn this switch on and the storage capacitor C will charge to the instantaneous value of the input voltage with a low time constant:

$$\tau_{CHG} = \left[R_0 + r_{DS(on)}\right] \cdot C \tag{9.58}$$

where $R_0$ is the very small output impedance of the voltage follower with $OA_1$ and $r_{DS(\text{on})}$ is the ON resistance of the FET.

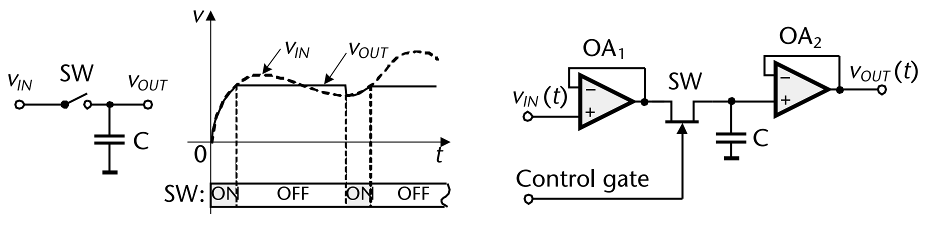
In absence of a positive pulse, the switch SW is turned off and the capacitor is isolated from any loads through the second voltage follower. For larger C values an isolation resistor (about 10 kΩ) should be included between C and the noninverting terminal of $OA_2$ to protect its input.

In order to prevent leakage of C, the $OA_2$ should be an op-amp with high impedance, (e.g., from TL08-series (op-amps with input JFET, manufacturer Texas Instruments)). The capacitor should be of high dielectric quality and free from polarization phenomenon or dielectric absorption [1]. An alternative option is to replace the $OA_2$ by a voltage follower with JFET and to close a negative feedback loop from follower output to the inverting input of $OA_1$. The feedback reduces the nonlinearity errors.

The electronic switch is imperfect and some input signal could appear in the output even in hold mode. This is called feedthrough, specified in decibels.

In the real word, a difference between the input and output is unfortunately available, which is called output offset and expressed in millivolts.

Two additional time factors influence the circuit operation: (1) the *aperture time* (typically below 0.1 $\mu$s), which is the delay between the time that the pulse is applied



(a) (b)

**Figure 9.20** Sample-and-hold circuit: (a) schematic, and (b) practical implementation.

to the SW and the actual time the switch closes, and (2) acquisition time, which is the time it takes for C to charge from the previous level of holding voltage to the new value after SW has closed. Aperture time is an attribute of SW, whereas the acquisition time could be reduced choosing an $OA_1$ with high output current capability.

The capacitor C creates an additional pole in the circuit transfer function, which must be reckoned. The pole limits the maximum frequency the S/H can acquire. This is called the full power bandwidth, specified in kilohertz or megahertz.

The major S/H parameter representing its ability to maintain the output in hold mode is featured by the droop rate, which is the amount of drift per time unit (e.g., 1 mV/s).

This feature is improved in a S/H circuit with infinite hold interval time, which is a mixed-mode schematic containing an analog input circuit and comparator as well as a digital loop: an enable-gate, a counter, and a digital-to-analog convertor (DAC) (Figure 9.21). The absolute value of the input is permanently compared with the absolute value of the output. If the input exceeds the output and the gate is enabled, in the next sampling time, clock pulses provide an increase of the counter digital value, which causes the step-by-step increase of the output as long as a new equality is established. Now, the counter stops and the new value is held in digital form (counter output) or analog (DAC output).

S/H circuits are available on a single monolithic chip, such as the series LFx98x (manufacturer is Texas Instruments) [16]. The device has two inputs for the signal to be sampled and for the hold pulse. The capacitor should be externally added.

## 9.4 Signal Linearization, Linearization Methods

Nonlinearity is present in any physical device for different reasons and at varying degrees.

The global nonlinearity of a sensor, already defined in Section 1.3.3 and Figure 1.5, shows a segmentation of the total nonlinearity error in three parts:

1. Offset error;
2. Sensitivity error;
3. Linearity error.

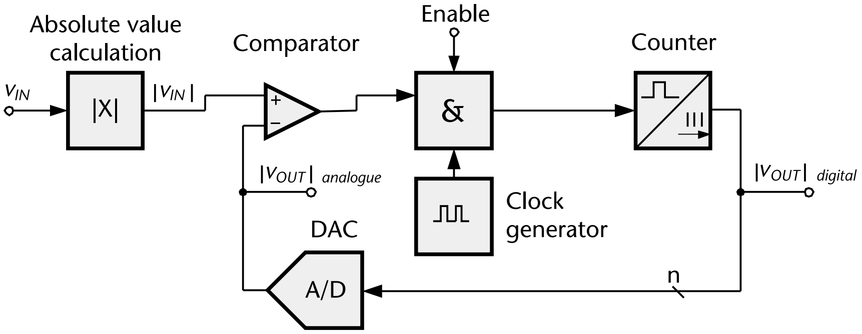


Figure 9.21 Mixed-mode S/H circuit with infinite hold interval time.

If the level of total nonlinearity can be neglected without exceeding the error tolerances (best case), the sensor can be assumed linear and can be expressed by a linear analytical model. Furthermore, the principle of superposition holds for linear systems. If nonlinearities are exhibited in the static input/output (I/O) operating curve, the overall performance of the sensor application will be degraded.

The offset and sensitivity errors are theoretically cumulative and multiplicative errors. It is relatively easy to determine and compensate for them.

If the linearity error is not negligible, its reduction or elimination is much more difficult and complex. This section will present improvement methods, which are also suitable for inductive sensors.

Nonlinearities of a sensor can manifest in a static form (where the output is not proportional to the input) or—more inconveniently—in a dynamic form such as the jump phenomena or frequency instability.

The sensor static I/O curve is generally monotonic and the improvement of the linearity makes sense. A classical linearization process must compensate for the static nonlinear relationship between the sensor input and its output.

If the I/O curve unfortunately has nonmonotonic regions including saturation, hysteresis, or dead zone, classical linearization is not applicable.

Linearization of the static I/O curve should begin with immanent measurements: a suitable choice of the sensing principle and signal processing (e.g., differential methods), with adequate design of the sensing element and sensor electronics, and so forth.

If these measures are applied but the I/O curve still requires further amelioration, then the option is to use linearizing by modifying the output itself, therefore using a calibration and linearization process. This basically consists of two steps (Figure 9.22):

1. A zero and span correction performed in the primary sensor or in an additional circuitry.

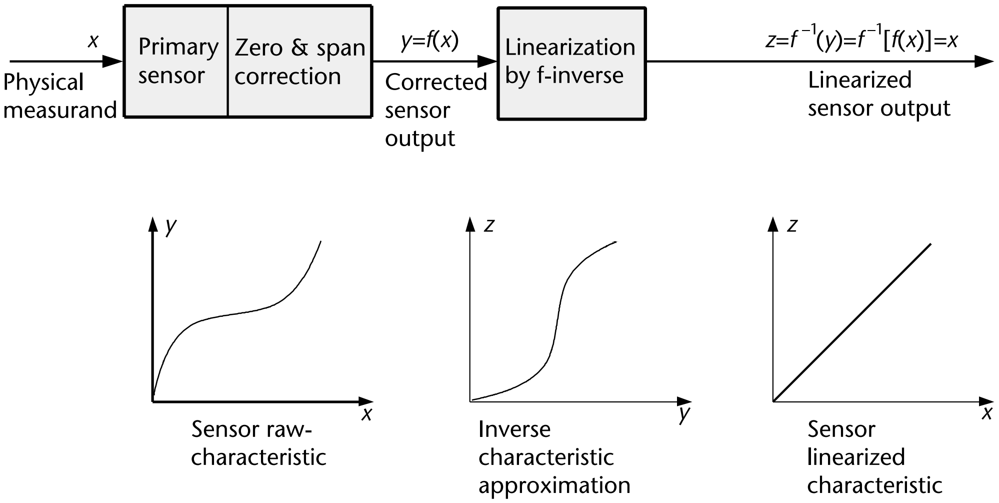

x
Physical measurand
Primary sensor
Zero & span correction
y=f(x)
Corrected sensor output
Linearization by f-inverse
z=f^{-1}(y)=f^{-1}[f(x)]=x
Linearized sensor output
y
x
Sensor raw-characteristic
z
y
Inverse characteristic approximation
z
x
Sensor linearized characteristic

**Figure 9.22** Basic representation of the calibration and linearization process in a sensor [17].

2. A linearization, which is carried out in a specific linearization stage. If the I/O curve is given by the nonlinear function $y = f(x)$, then the relationship between the input $y$ and the output $z$ of the linearization stage should be $f^{-1}$, so that:

$$z = f^{-1}(y) = f^{-1}[f(x)] = x \tag{9.59}$$

where $f^{-1}$ is the inverse transfer function.

The correction and linearization process is usually inserted into the sensor signal processing chain (see the stages SCD and LIN in the generic structure of the inductive sensor evaluation electronics in Figure 7.1).

Static correction and linearization of the sensor can be made in several ways as follows: using analog hardware, using digital hardware, using nonlinear calibration curves, nonlinear transformations, and associated software, and using artificial intelligence.

### 9.4.1 Analog Hardware-Based Linearization

In the analog hardware approach, the analog linearizing circuit is directly connected to the primary sensor output and ensures that its output is proportional to the input of the original device (Figure 9.22).

Three types of correction and linearizing hardware should be mentioned:

1. Offsetting circuitry;
2. Circuitry providing a proportional output;
3. Curve shapers.

An offset error (see Figure 1.5) is a nonlinearity that can be easily removed by analog devices. This is accomplished by adding a DC quantity of equal value to the offset error in the opposite direction and is called offset compensation.

Beside this removal of the offset, the case of deliberate addition of an offset, called offsetting, has a large application field in the control and measuring devices.

Offset compensation or offsetting can be implemented in different sensor stages using fine adjustment compensating resistors or higher feedback reverse transmission in negative feedback loops (see parameter $\beta$ in Figure 8.9). The most efficient method is the using of a differential amplifier or a summing amplifier (e.g., Figure 9.2(a)).

The proportional-output hardware is implemented in reactive sensors in conjunction with the bridge circuits. The relation between the bridge output and the change in resistance of the sensing element provoked by the measurand is generally nonlinear and depends on the bridge type: constant-voltage, constant-current and constant-voltage half-bridge. The constant-voltage bridge is characterized by a medium nonlinearity in comparison with the lowest nonlinearity (by constant-current bridge) and the highest nonlinearity (by constant-voltage half-bridge).

A possibility to improve the linearity of the most used constant-voltage bridge is to replace the traditional Wheatstone bridge by a proportional-output circuit with an op-amp (Figure 9.23). Its transfer function:

$$v_{\mathrm{OUT}} = \frac{\left(R_2R_3 - R_1R_4\right)}{R_3\left(R_2 + R_4\right)} \cdot v_{REF} \tag{9.60}$$

is slightly different from the Wheatstone bridge transfer function:

$$v_{\mathrm{OUT}} = \frac{\left(R_1R_4 - R_2R_3\right)}{\left(R_1 + R_2\right)\left(R_3 + R_4\right)} \cdot v_{REF} \tag{9.61}$$

Let us suppose that both circuits are balanced; that is, all resistances are equal to a value $R$. A variation $\delta R$ of the active resistor due to the measurand change causes a nonlinear change $\delta v_{\mathrm{OUT}}$ of the bridge output. In contrast, according to (9.60), the change $\delta v_{\mathrm{OUT}}$ at the op-amp is proportional to $\delta R$ (linearity):

$$\frac{\delta v_{\mathrm{OUT}}}{v_{REF}} = -\frac{1}{2} \cdot \frac{\delta R}{R} \tag{9.62}$$

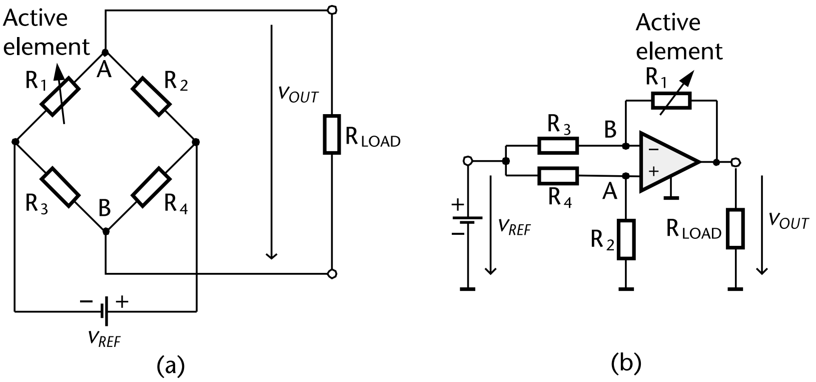
and is two times higher than the best case change of the bridge output (double sensitivity).

Curve shapers are real linearity improvers and are circuits with allocable transfer characteristics. The conventional means are schematics, which accomplish the inverse transfer function *f*-inverse of the sensor (Figure 9.22).

In the simplest implementation, curve shapers are amplifiers with adjustable gain. A typical arrangement is the inverting amplifier (see Table 9.3).

The feedback resistance $R_R$ is adjustable by some means:

- A bank of resistors connected in parallel through solid-state switches, which are controlled by an ADC convertor;
- A bank of resistors connected in parallel by automatic switching realized by using Zener diodes, which start conducting at certain voltage levels.



(a) (b)

**Figure 9.23** The replacement of the traditional Wheatstone bridge (a) with a proportional-output circuit and (b) with op-amp.

In both cases, the gain is variable in discrete steps:

- Nonlinear amplifier with diodes or bipolar transistors in the feedback loop;
- Potentiometers, which are manually or automatically adjusted.

In these cases, the gain can be continuously adjusted.

### 9.4.2 Software-Based Linearization

In the software approach to linearization, the output of the sensor (signal $y$ in Figure 9.22) is read into a digital processor with software-programmable memory and the output is modified according to the program instructions to get a linear global I/O relation. If the nonlinear relationship between the input and the output of the device to be linearized is known, the input can be computed for a known value of the output by means of a processor and its memory.

Three approaches can be used:

1. Variable transformation;
2. Equation inversion;
3. Look-up table (LUT).

In the first method, the nonlinear characteristic of the original sensor is known in the analytic form (i.e., an algebraic model):

$$y = f(x) \tag{9.63}$$

where $x$ is the sensor input and $y$ the sensor output (Figure 9.22).

The variables $x$ and $y$ are transformed into two new variables $x'$ and $y'$ (e.g., log of variable) so that the relationship between $x'$ and $y'$ is linear. Then, linear methodologies can be used to analyze the transformed data of $x'$ and $y'$ without sacrificing any accuracy. Since the transformation relation is known, the results can be transformed back into the domain of variables $x$ and $y$.

In the second method, assuming that (9.63) is a one-to-one relationship, a unique inverse function can be determined, as given by the following equation:

$$x = f^{-1}(y) \tag{9.64}$$

This equation may be programmed as a computation algorithm into a nonvolatile microcontroller memory. When an output value $y$ is measured and supplied to the controller, the processor will compute the corresponding input value $x$ and provide this as the final output.

Characteristic for the first two methods is the fact that they require a priori knowledge of the sensor characteristic (9.63). This requirement can be fulfilled applying a curve fitting method or a progressive polynomial calibration [17].

For the curve fitting, it is recommended to perform the linearization in the digital domain, which means taking the following steps:

- Analog-to-digital conversion of the sensor output $y$;
- Digital curve fitting in an integrated microcontroller with memory facility. Typically, this is done by a calibration process where a number of $x$,$y$-pairs are experimentally registered and the dependence (9.63) is determined by off-line curve fitting.

  This can be achieved by linear or polynomial approximation with a polynomial of degree $n$, equal to the number of measured sensor outputs (usually $n < 6$), using for example the MATLAB commands *nlinfit()* or *polyfit()*.

  When the sensor output is measured for a too-small set of known input signals an interpolation algorithm is then used to compute the sensor transfer function (9.63).

  As soon as the inverse transfer function is calculated, the input signal can be derived from the sensor output [see (9.59)].
- Digital-to-analog conversion to generate the analog linearized sensor output.

The advantages of curve fitting are that an inverse transfer function is obtained for a complete signal range with low memory requirements. The disadvantage can be that sometimes higher-order polynomial are necessary to get the desired accuracy. This involves more complex computation and memory capacity.

Progressive polynomial calibration operates based on the principle that each calibration measurement is directly used to calculate one calibration coefficient in the correction function (offset, gain, and nonlinearity correction). The correction is applied to the sensor output and each step is independent of the previous one.

The benefits are that it gives good linearization for a minimum number of points with resulting low memory requirement and uses a repetitive algorithm (step-by-step calibration).

Linearizing sensor output is one of the major application of the look-up table procedure.

If the sensor has a powerful controller (with 32-bit floating-point processing, hardware multiplication, etc.), then the inverse function $f^{-1}$ can be calculated in real time and LUT is unnecessary.

Real inductive sensors are compact and low-cost mass products and contain at their best a small embedded 8-bit system with limited computational power. For such systems, LUT is very suitable.

Effectively, a sufficiently large number of pairs of value ($y$, $x$) are stored in the memory of a microcontroller in form of a table of ordered pairs [18]. These values should cover the entire sensor operating range.

In hardware form, $f^{-1}$ is typically implemented as a LUT in flash memory. Usually $y$ is the $n$-bit integer produced by the ADC and this integer is used as a pointer to the memory location where $f^{-1}$ ($y$) with $m$-bit resolution is stored (Figure 9.24).

When a value for $y$ is entered into controller, the processor scans the stored data to check whether that value is present. If so, the corresponding value of $z = x$ will be read, and this is the linearized output. If the value of $x$ is not present (due to the limitation of memory capacity), the processor will interpolate the data in the vicinity of this value and will compute an approximated output. In the linear interpolation method, the neighborhood of the data table where the $y$ value falls is fitted with a straight line, and the $x$ value is computed using this straight line.

High-order interpolations use nonlinear interpolation curves such as quadratic and cubic polynomial equations (splines) [19].

With respect to the accuracy, the variable transformation and equation inversion are more accurate than the LUT method, which has a limited granulation.

Regarding the speed, they are relatively slow because the data is transferred, transformed, and processed using program instructions, which are stored in the program memory and are accessed in a sequential manner. The LUT is faster while it uses a parallel data access.

LUT is a memory-intensive method. The better the specified accuracy, the higher the table (i.e., the memory capacity). Since the entire data table has to be scanned to check for a given data value, the accuracy increase is derived at the expense of speed. The first two methods do not need excessive memory for data storage.

### 9.4.3 Logic Hardware Linearization

In the logic hardware approach to linearization, the output of the sensor (signal $y$ in Figure 9.22) is read by a device with fixed logic circuitry for processing the data. The resulting output is also digital.

The software linearization is flexible because the algorithm ca be simply modified (improved, changed) by altering the program stored in the controller memory. In addition, complex nonlinearities can be handled. However, the method is relatively slow.

In the logic hardware method, the linearization algorithm is permanently implemented in the IC form using appropriate digital logic circuits for data processing and memory elements (wired logic). The algorithm and numerical values of the model parameters cannot be modified without redesigning the IC chip. These chips are mass-produced for an extensive commercial market. They do not include complex algorithms, and consequently, offer moderate results but they have accessible prices. Since both the access of stored program instructions and the extensive data manipulation are not involved, the hardware method of linearization is substantially faster than the software method.

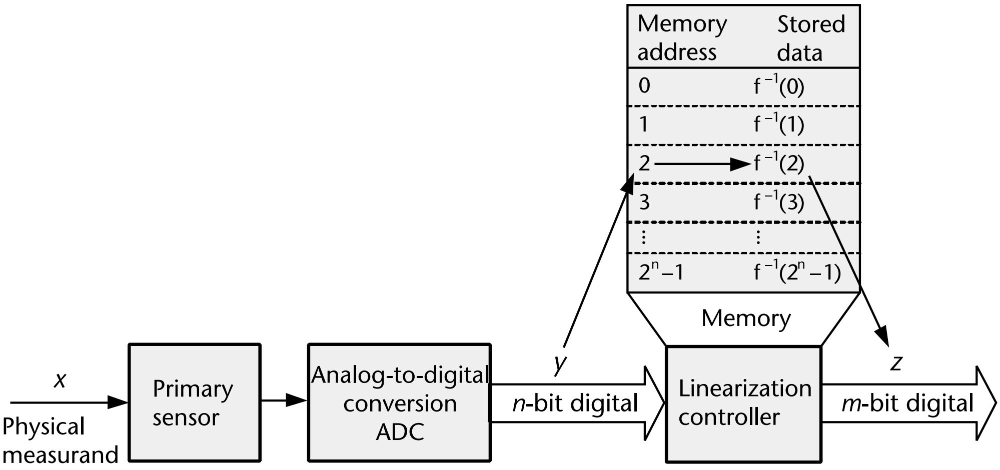


**Figure 9.24** Linearizing by LUT.

### 9.4.4 Hardware-Software Mixed Approaches

Currently, mixed software-hardware methods are promoted. An integrated microcontroller unit with processor and EEPROM memory, whose content can be modified, is a hardware option that benefits the software procedure. Another option is the implementation of the modern systems as field-programmable gate arrays (FPGA), (manufacturers such as XILINX, ALTERA), programmable system on chip (PSoC) (Cypress Semiconductor).

*Example 9.3*

The next description aims to show the concrete results obtained by the shaper linearization of an already analyzed oscillator for inductive sensors.

Let us recall the pilot ring-circuit oscillator with LC dipole (Section 8.3.2, Figure 8.17). Its output signal, expressed by (8.80):

$$v_{LC}(t) = \frac{4 \cdot V_{SD}}{\pi} \cdot \frac{R_P}{R_{ADJ}} \cdot \sin\left(2\pi f_0 t + \varphi\right) \tag{9.65}$$

has an amplitude that is proportional to the equivalent loss resistance $R_P$ of the resonant LC circuit (see (4.38)):

$$V_{lcm} = \frac{4V_{SD}}{\pi R_{ADJ}} \cdot R_P \tag{9.66}$$

and this resistance has a dependence on the distance $s$ between the target and the inductive sensing element, which can be favorably approximated by an exponential function (see (8.82)):

$$R_P(s) = R_{P\infty}\left[1 - e^{\frac{-(s-s_0)}{SK}}\right] = R_{P\infty} - R_{P\infty} \cdot e^{\frac{-(s-s_0)}{SK}} \tag{9.67}$$

For the definition of the system parameters: $R_{P\infty}$, $SK$, and $s_0$, see (8.82).

Figure 8.18, which shows the form of this curve (for parameters: $R_{P\infty} = 27.35$ kΩ, $SK = 1.4$ mm, and $s_0 \approx 0$ mm), predicts a successful application of the analogue hardware-based linearization.

After the substitution of (9.67) into (9.66) and the elimination of the constant term corresponding to $R_{P\infty}$ (offset correction), the sensor raw-characteristic results in:

$$V_{lcm}(s) = \frac{4V_{SD}R_{P\infty}}{\pi R_{ADJ}} \cdot e^{-\frac{s}{SK}} \tag{9.68}$$

The signal-flow graph and the block schematic of the linearization path are shown in Figure 9.25(a). Due to the natural-exponential dependence in (9.68), the inverse transfer function results in a natural logarithm: $f^{-1} = \ln y$.

Figure 9.25(b) illustrates a quick-and-dirty method to graphically check the cascade-structure graph in Figure 9.25(a). The first quadrant contains the raw

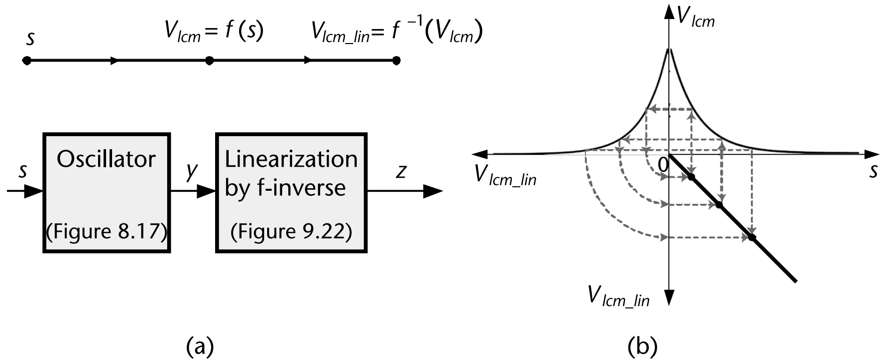

s
$V_{lcm} = f(s)$
$V_{lcm\_lin} = f^{-1}(V_{lcm})$
s
Oscillator
(Figure 8.17)
y
Linearization by f-inverse
(Figure 9.22)
z
$V_{lcm}$
$V_{lcm\_lin}$
0
s
$V_{lcm\_lin}$
(a)
(b)

**Figure 9.25** Concrete embodiment of the general architecture in Figure 9.22 for a ring-circuit oscillator with LC dipole: (a) Block diagram and associated signal graph, and (b) graphical explaination.

characteristic (9.68) and the second quadrant contains the inverse function (natural logarithm with the vertical axis used as input axis).

To obtain the reference points of the global characteristic (in the fourth quadrant), simply

- Consider a starting value of *s* and draw a vertical auxiliary line until the intersection with the curve in the first quadrant;
- Continue with a horizontal line from this intersection point until the intersection with the curve in the second quadrant;
- Continue with a vertical line until the intersection with the negative abscissa axis;
- Ship this point on the negative ordinate axis making a 90-degree circle arrow;
- Having the coordinates, represent the first reference point in the fourth quadrant;
- Repeat the steps above for more reference points and fit the straight line.

This linearizing method was applied using the circuit in Figure 9.26. The schematic consists of

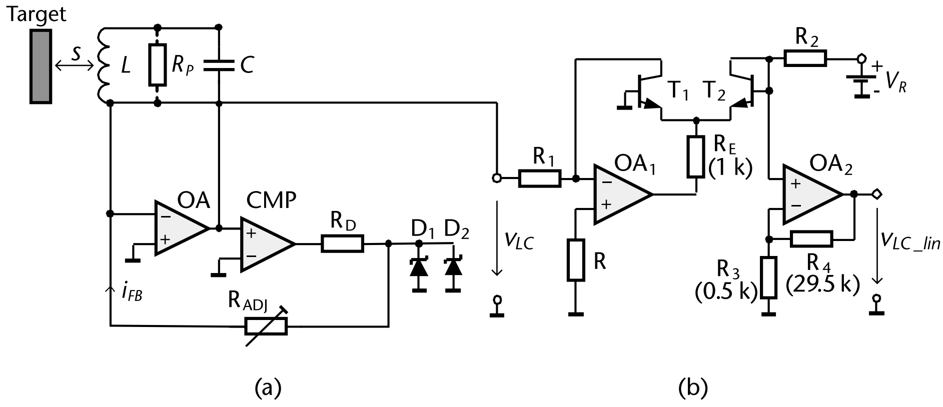

Target
s
L
$R_P$
C
OA
CMP
$R_D$
$D_1$ $D_2$
$i_{FB}$
$R_{ADJ}$
$V_{LC}$
$R_1$
$OA_1$
R
$T_1$ $T_2$
$R_E$
(1 k)
$R_2$
+
-
$V_R$
$OA_2$
$R_3$
(0.5 k)
$R_4$
(29.5 k)
$V_{LC\_lin}$
(a)
(b)

**Figure 9.26** Circuit diagrams of the ring-circuit oscillator with LC dipole (a) followed by the linearizing stage with a logarithmic amplifier (b) with two op-amps and two transistors ($T_1$ and $T_2$).

a. The above-mentioned pilot oscillator (see Figure 8.17);
b. A nonlinear amplifier with two op-amps and two bipolar transistors, which has a natural-logarithmic transfer function.

Considering an inverting amplifier with op-amp (see Table 9.3), the replacement of the feedback resistor by a diode leads to an elementary logarithmic amplifier having a nonlinear transfer function due to the nonlinear v-to-i diode characteristic. To a larger dynamic range, the diode can be replaced by a bipolar transistor, which also has a nonlinear dependence $i_C$ versus $v_{BE}$ (see (8.110)). The temperature dependence due to the temperature voltage $V_T$ can be improved using a differential amplifier with two matched transistors (Figure 9.26(b)).

The transfer function of this amplifier is given by:

$$V_{lcm\_lin}(s) = -V_T \frac{R_3 + R_4}{R_3} \cdot \ln\left[\frac{R_2}{R_1} \cdot \frac{V_{lcm}(s)}{V_R}\right] \tag{9.69}$$

where $V_R$ is a scale-adjustment reference voltage. The high-performance schematic satisfies the (9.69) over a dynamic range of four decades, from input voltages of 2 mV to 20V.

At the end, Figure 9.27 shows the experimentally obtained curve: $R_P$ as a function of the distance $s$ and the resulting graph: $V_{lcm\_lin}$ versus $s$ after linearization (right-side ordinate axis).

### 9.4.5 Artificial Neural Networks Approaches

A forward-looking method for linearizing the response of a sensor is the use of artificial neural networks (ANNs) taking advantage of their nonlinear modeling and learning from examples capabilities, and without requiring that the sensor model be available. Advantageously, the method requires embedded systems based on simple, low-resolution microcontrollers (8- or 16-bit controller) or fixed-point DSPs.

The ANN is trained with the following input-output data: (a) the nonlinear output of the sensor is used as input data and (b) the difference between sensor outputs and the desired linear responses are target values. The ANN is programmed into the system memory; thus, the solution does not require additional hardware.

Reference [20] presents a case study for the linearization of a temperature sensor with NTC resistor. This is a general procedure and could also be applied for inductive sensor linearization.

The small used 1-2-1 ANN is a multilayer perceptron (MLP) and contains a minimal set of four neurons (processing elements): one input neuron, two hidden ones, and one output, as well as their seven connections. Each connection is modeled by a weight.

The ANN learns (adjust its weights) from input-output data (examples) and this consists of training MLP with experimental data measured to provide the desired linear output. According to the results of [20], simple MLPs are able to linearize a broad class of different nonlinear sensors. In addition, ANN extends the linear response range without supplementary electronics.

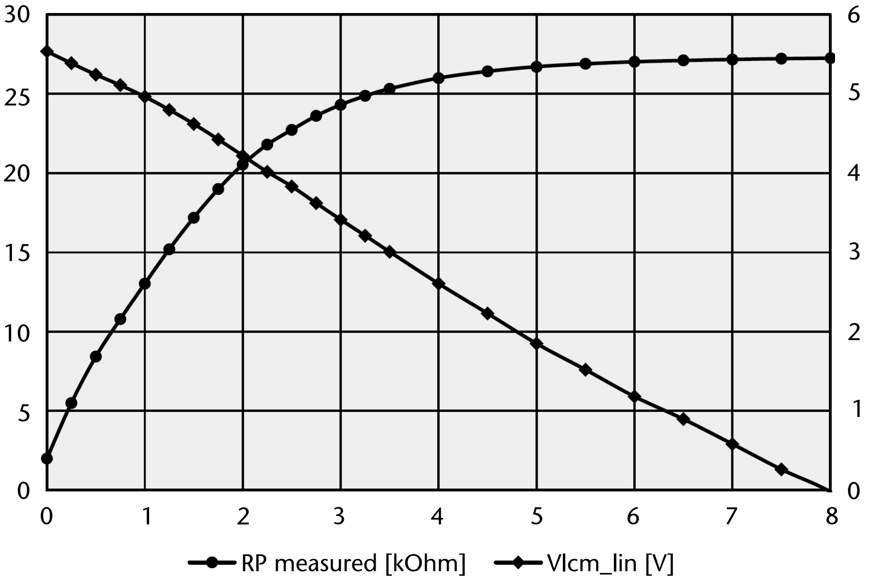


**Figure 9.27** Experimental results of the linearization process performed in the circuit in Figure 9.26. *X*-axis is the distance s [mm] axis, and the coordinate for $V_{icm}$ is the right *y*-axis.

The ANN implementation can be summarized by the following steps: (a) ANN training in a computer to determine the weight values, (b) values quantization with the wanted precision, and (c) final hardware implementation into a controller or DSP (LUT or piecewise linear approximation of the nonlinear neurons output function).

## 9.5 Comparators, Window Discriminators

Voltage comparators have two inputs, $v_{\mathrm{IN+}}$ and $v_{\mathrm{IN-}}$, and one output, $v_{\mathrm{OUT}}$. An op-amp operated open-loop becomes a comparator (Figure 9.1(a)).

Comparators are used to compare an input time-varying signal applied at the noninverting/inverting input with a fixed or variable reference voltage existing at the inverting/noninverting input.

Accordingly, the circuit operates as noninverting or as inverting comparator. Its output changes from a logical 0 [$v_{\mathrm{OUT}} = \mathrm{V}(0)$] when the input is below the reference to a logical 1 [$v_{\mathrm{OUT}} = \mathrm{V}(1)$] once the input signal exceeds the reference (or vice versa).

This amplitude-discrimination property may be used to quantize analog signals in an analog-to-digital converter or to discriminate a pulse height. By setting the reference level at 0V, the comparator functions as a sense or polarity detector.

In contrast to an ideal comparator, the practical circuit has a total input swing $\Delta v_{\mathrm{IN}}$ that is nonzero and should be as low as possible.

The basic voltage comparator uses a differential-amplifier transistor pair where the input signal, applied at one input, is compared with the reference applied at the other input. Such circuits were presented and described in Figure 8.24 (bipolar transistors $T_1$ and $T_2$, assisted by a current sink with $T_3$ and $T_4$ as the common-emitter load), and Figure 8.25 (MOS transistors $M_1$ and $M_2$ connected to a current source with $M_3$ and $M_4$ as a common-source circuit).

Theoretically, this elementary stage has a wide transition region ($\Delta v_{IN} = 4.5\ V_T = 117$ mV, typically), which can be essentially reduced by cascading this stage with additional high-gain stages and using only a small portion of the transition region.

Since a comparator is not intended to be used with negative feedback, frequency-compensation components may be omitted, so that greater bandwidth (high speed) is attainable compared to op-amps.

The circuits can be realized discrete or in integrated form. For example, the legendary comparator LM710 (manufacturer former National Semiconductor, now Texas Instruments) has a typical gain of 1,500 and $\Delta v_{IN} = 2$ mV. Up-to-date circuits of the series LM111, 211 and 311 have similar features. In addition, they contain an open-emitter and open-collector output switching transistor [21]. Packages with two or four independent comparators are also available.

The voltages V(0) and V(1) are compatible with various logic levels (TTL, MOS, etc.) or are indefinite. To obtain defined output voltages, which are independent of the supply voltages, two back-to-back Zener diodes and a current-limiting resistor can be added to clamp the comparator output (see the clamping with Schottky barrier diodes in Figure 8.17(a)).

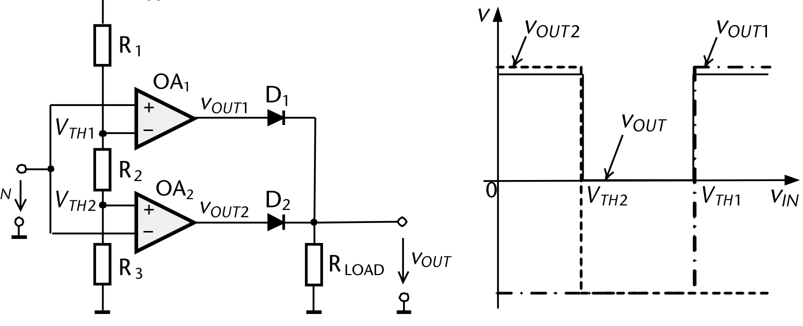
An important comparator application is the window discriminator. The circuit compares an input voltage with two already-set threshold voltages and signalizes by the level of its output if the input signal lies in the voltage window or outside of this range.

The schematic consists of two comparators, which compare the input voltage $v_{IN}$ with the thresholds $V_{TH1}$ and $V_{TH2}$ generated by the voltage divider $R_1$–$R_2$–$R_3$ (Figure 9.28(a)). Their expressions are:

$$V_{TH1} = \frac{R_2 + R_3}{R_1 + R_2 + R_3} \cdot V_{CC} \text{ and } V_{TH2} = \frac{R_3}{R_1 + R_2 + R_3} \cdot V_{CC} \tag{9.70}$$

To generate the composite output voltage $v_{OUT}$, the comparator outputs are linked in a logical OR-operator (diodes $D_1$ and $D_2$).

If $v_{IN}$ lies in the range between the thresholds, the comparator outputs are LOW, the diodes are turned off and $v_{OUT} = 0$V. When the input exceeds the high threshold



(a) (b)

**Figure 9.28** Window discriminator: (a) possible implementation with switching diodes, and (b) waveforms of the outputs and composite input-output curve (solid line).

$V_{TH1}$ or goes below the low threshold $V_{TH2}$, the output of $OA_1$ or $OA_2$ goes HIGH, turns ON the corresponding diode and clamps $v_{OUT}$ on its HIGH level.

The windows comparator can be realized with op-amps but it is also available in IC form, for example, in TCA965B (manufacturer Infineon). A similar function is integrated in the circuit 555 (see Figure 8.44).

## 9.6 Regenerative Comparators (Schmitt Trigger)

In the case of a conventional comparator open-loop-operated, when noisy inputs approach the transition region, the output could suffer jitter effects evident in fast consecutive output changes between HIGH and LOW levels. This undecided behavior can be suppressed implementing a positive feedback (Figure 9.29(a)).

Applying the positive feedback to the noninverting input of a classical comparator circuit, a regenerative comparator (with hysteresis) or Schmitt trigger (SCT) results. This active circuit, which converts the analog input signal to a digital output signal is called a trigger because the output retains its value until the input changes sufficiently to trigger a change.

Here too, discrete SCT implementation was used in the past. The book presented two versions in Figure 7.3. Presently, SCTs are realized with integrated op-amps (Figure 7.15) or comparators.

In the noninverting configuration, when the input is higher than a chosen threshold, the output is HIGH. When the input is below a different chosen threshold the output is LOW, and when the input is between the two levels the output retains its value.

This dual threshold action is called hysteresis and implies that the SCT possesses memory and can act as a bistable multivibrator (latch or flip-flop). The major advantage of the regenerative comparator is the inherent self-latching action of the hysteresis. There is a narrow relation between the two kinds of circuits: a SCT can be converted into a latch and a latch can be converted into a SCT.

Due to the hysteresis, which is larger than the transition region of a usual comparator, the switching error is larger but the decision is firm.

A research of several technical publications led to the following popular Schmitt trigger versions:

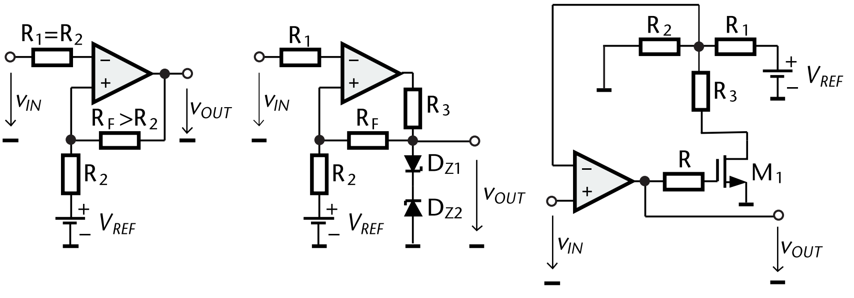
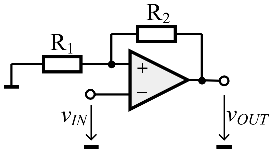
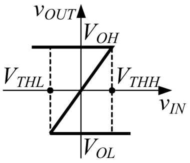
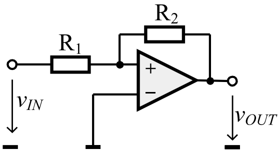
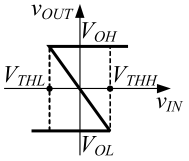


**Figure 9.29** Schmitt trigger circuits: (a) traditional version, (b) version with clamping Zener diodes, and (c) precision version with large hysteresis.

**Table 9.5** Schmitt Trigger with Op-Amps

| *Circuit Type* | *Schematic* | *Transfer Characteristic* | *Threshold Expressions* |
|---|---|---|---|
| Inverting trigger |  |  | $V_{THH} = [R_1/(R_1 + R_2)] \cdot V_{OH}$<br>$V_{THL} = [R_1/(R_1 + R_2)] \cdot V_{OL}$ |
| Noninverting trigger |  |  | $V_{THH} = (R_1/R_2) \cdot V_{OH}$<br>$V_{THL} = (R_1/R_2) \cdot V_{OL}$ |

1. STG without firmly defined output levels;
2. STG with defined output levels;
3. Precision STG with large hysteresis.

The schematic of the first category is shown in Figure 9.29(a). The existence of a nonzero DC reference voltage $V_{REF}$ is a requirement.

The operation can be followed in the graphs of Figure 9.30. Starting with $v_{IN} = 0$, the op-amp output goes to the highest value $+V_0$, which is smaller than the positive supply voltage by the output saturation voltage (Figure 9.30(a)). Consequently, the threshold voltage at the noninverting terminal takes its high value $V_{THH}$:

$$V_{THH} = \frac{R_F}{R_F + R_2} \cdot V_{REF} + \frac{R_2}{R_F + R_2} \cdot V_O \tag{9.71}$$

When $v_{IN}$ reaches this threshold, the op-amp switches over, its output goes on the lowest level $-V_0$ and the threshold changes on its low value $V_{THL}$:

$$V_{THL} = \frac{R_F}{R_F + R_2} \cdot V_{REF} - \frac{R_2}{R_F + R_2} \cdot V_O \tag{9.72}$$

Let us suppose now a return to zero of the input voltage starting from high values. If this voltage achieves the threshold $V_{THL}$ (Figure 9.30(b)) a switching process of the op-amp output from $-V_0$ to $+V_0$ occurs and the operation cycle is closed. The final result consisting of a switching curve with hysteresis is shown in (Figure 9.30(c)).

The circuit allows an independent adjustment of the parameters: the hysteresis middle point by $V_{REF}$ and the width by the levels $-V_0$ and $+V_0$, which usually have equal absolute values.

Unfortunately, these levels vary with the load at the amplifier output, temperature, and so forth, and thus, the switches thresholds suffer undesirable fluctuations.

To prevent this drift, a simple method is to clamp the levels by means of an antiserial network containing two identical Zener diodes (Figure 9.29(b)). For every polarity of the op-amp output, one diode is forward-biased and the other reversed-biased. Hence, the levels $-V_0$ and $+V_0$ have defined equal absolute values (one Zener voltage plus one silicon diode voltage). Thus, the thresholds $V_{THL}$ and $V_{THH}$ are now constant.

Absolutely stable thresholds can be ensured with the circuit in Figure 9.29(c), which contains a network $R_1$ to $R_3$ and a switch realized with the MOS transistor $M_1$. The direct derivation of the thresholds from the voltage reference $V_{REF}$ guarantees the absolute stability of these.

Starting with $v_{IN} = 0$, the op-amp goes LOW, $M_1$ is switched off and the voltage divider consists of $R_1$ and $R_2$. Thus, the threshold voltage at the inverting terminal is:

$$V_{THH} = \frac{R_2}{R_1 + R_2} \cdot V_{REF} \tag{9.73}$$

When $v_{IN}$ reaches this threshold, the op-amp switches over, its output goes HIGH and $M_1$ is switched on. The resistor $R_3$ appears connected in parallel to $R_2$ and the threshold drops on the low value:

$$V_{THL} = \frac{R_2 R_3}{R_1 R_2 + R_1 R_3 + R_2 R_3} \cdot V_{REF} \tag{9.74}$$

Supposing now a return to zero of the input voltage starting from high values, if this voltage achieves the threshold $V_{THL}$, a switching process of the op-amp output from HIGH to LOW occurs and the standard operation cycle is closed.

Beside the stability, the circuit offers a large adjustment range of the thresholds.

The Schmitt trigger is frequently used in inductive sensors due to its important feature, being able to convert slowly varying weak sensor signals into a waveform, which displays an abrupt, almost discontinuous, change that occurs at a precise value of the sensor signal.

## 9.7 Phase-Locked Loop Circuits

A phase-locked loop (PLL) is a particularly flexible circuit building block (control system) that generates an output signal whose phase is related to the phase of an input signal [10, 22].

The elementary PLL consists of a variable frequency oscillator (VCO) and a phase detector (PD) in a feedback loop. The VCO generates a periodic signal, and the phase detector compares the phase of that signal with the phase of the input periodic signal, adjusting the oscillator to keep the phases matched.

Fundamentally, the PLL is not restricted to carrier phase tracking. Keeping the input and output phase in lock step also implies keeping the input and output frequencies the same. Consequently, the PLL can track an input frequency, or it can generate a frequency that is a multiple of the input frequency.

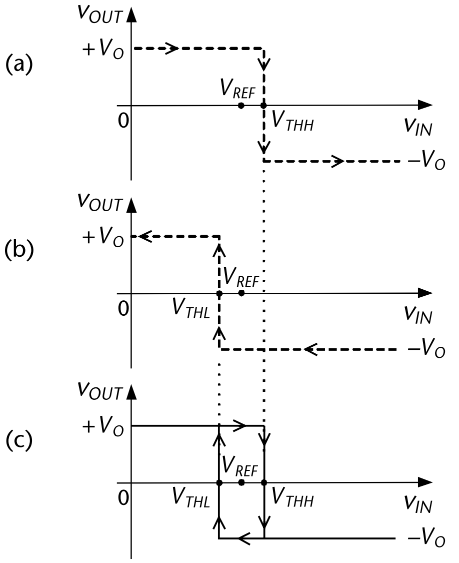


**Figure 9.30** Transfer characteristics of the inverting Schmitt trigger (see Figure 9.29(a)): (a) for increasing $v_{IN}$, (b) for decreasing $v_{IN}$, and (c) the composite input-output curve with hysteresis.

Moreover, when it is used to track other waveform attributes, such as sampling instant under the control of an interpolator, or the delay time of a locally generated waveform, the PLL concept still holds (Figure 9.31). The error detector can be very simple or complex and has to provide an S-shaped characteristic, which can be used under feedback control to drive the tracking error $e(t)$ or $e(n)$ between input waveform attribute $\varphi_r$ and generator waveform attribute $\varphi_o$ to zero. The characteristic expresses the dependence of $e(t)$ or $e(n)$ versus the difference of the attributes of the reference and tracked oscillator and has a linear form only in the vicinity of the origin.

Since a single IC can provide a complete PLL functionality, the technique is widely used for output frequencies from a fraction of a hertz up to many gigahertz. PLLs are often used in inductive sensors to demodulate a critical signal (recover a signal from a noisy source), to generate a stable frequency at multiples of an input frequency (frequency synthesis), or to provide precisely timed clock pulses for sensor digital logic circuits.

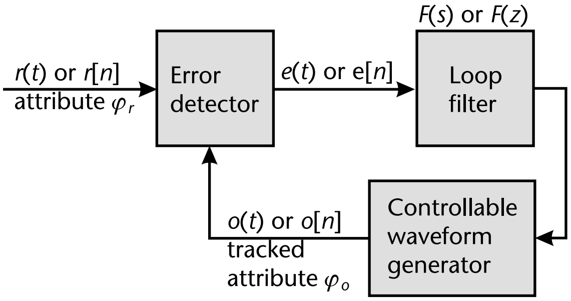


**Figure 9.31** High-level PLL block diagram.

### 9.7.1 PLL Concepts

The operation of a PLL is based around the idea of comparing the phase of two signals. This information about the phase difference between the two signals is then used to control the frequency of the loop.

As mentioned in the previous sections, a sinusoidal signal can be represented by a phasor, which rotates in the $x$-$y$ plane, drawing a circle if its amplitude is constant and having a certain angle speed; that is, angular frequency (velocity) $\omega$ (see (4.5)). In inductive sensors, this amplitude is often modulated and as time progresses, the phasor arrow oscillates above and below the circle shape, repeating itself after each cycle.

The beginning of the cycle can be represented as a particular point on the circle and as time progresses the point moves around the circle. Thus a complete cycle is equivalent to 360° or $2\pi$ radians. The instantaneous position on the circle represents the phase at that given moment relative to the beginning of the cycle.

Considering two signals with the same frequency, the peaks and troughs do not occur generally in the same place. There is said to be a phase difference between the two signals and this phase difference is measured as the angle between them. The most comfortable is the measurement between the zero crossing points.

When there are two signals that have different frequencies, it is found that the phase difference between the two signals always varies. The reason for this is that the time for each cycle is different, and accordingly, they are moving around the circle at different rates.

It can be inferred from this that the definition of two signals having exactly the same frequency is that the phase difference between them is constant. If the phase difference is fixed it means that one is lagging behind or leading the other signal by the same amount, constantly.

### 9.7.2 Analog PLL: Architecture and Operation

Although a PLL performs its actions on a voltage signal, PLL is basically a form of servo loop and all the basic criteria for loop stability and other parameters are the same [23].

The classical analog PLL (APLL) consists of three basic elements (Figure 9.32):

1. PD, usually an analog multiplier, which compares the phase of two signals (reference and VCO oscillator output) and generates a voltage according to the phase difference between these two signals.

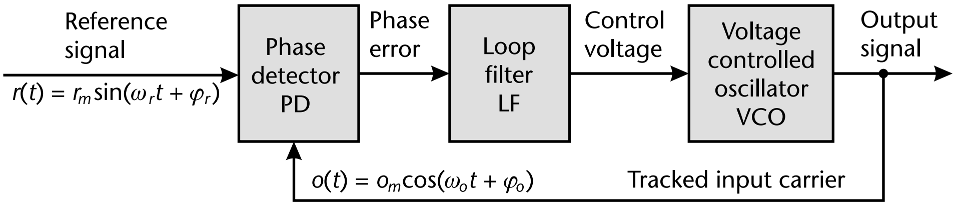

Reference signal
$r(t) = r_m \sin(\omega_r t + \varphi_r)$
Phase detector PD
Phase error
Loop filter LF
Control voltage
Voltage controlled oscillator VCO
Output signal
$o(t) = o_m \cos(\omega_o t + \varphi_o)$
Tracked input carrier

**Figure 9.32** Basic APLL for tracking the phase of a carrier signal.

2. Loop filter (LF), a low-pass filter (LPF) that is used to filter the output of the PD. It removes any high-frequency elements on the signal and also governs many of the characteristics of the loop and its stability.
3. VCO, which generates the PLL output signal, whose frequency can be controlled and swung over the operational frequency band of the loop. The idea is that the VCO will lock onto the reference signal and thus can be used to track a periodic signal as its phase and frequency varies.

The VCO produces an output signal that enters the PD. Here the phase of the signal $v_O$ from the VCO and of the incoming reference signal $v_R$ are compared and a resulting difference or error voltage is generated. This voltage corresponds to the phase difference between the two signals, it is passed through the LF, and then applied to the VCO as its tuning voltage.

Initially, the PLL loop will be out of lock and the error voltage will pull the frequency of the VCO toward that of the reference until it cannot reduce the error any further and the loop is locked. When the PLL is in lock a steady-state error voltage is produced.

By using an amplifier between the PD and the VCO, the actual error between the signals can be reduced to very small levels. However, some voltage must always be present at the control terminal of the VCO as this is what maintains the correct frequency.

The fact that a steady error voltage is present means that the phase difference between the reference signal and the VCO is not changing. As the phase between these two signals is not changing, it means that the two signals are on exactly the same frequency.

Again, if the frequency of the reference signal is equal to the frequency of the VCO, the phase difference is theoretically zero. The question is, what happens if one frequency drifts (usually of the reference signal). Two operation-specific PLL features respond to this question and have a great importance. They are the capture range $B_C$ and the tracking range $B_T$.

The tracking range $B_T$ is defined as maximal value of the difference $\omega_O - \omega_R$, whereas the equality $\omega_O = \omega_R$ can still be maintained. The capture range $B_C$ is defined as maximal value of the difference $\omega_O - \omega_R$, whereas the equality $\omega_O = \omega_R$ can be achieved. Although the features seem to be identical, the difference is essential: $B_T$ refers to a PLL, which is already locked; meanwhile, $B_C$ reflects the PLL ability to lock on a wanted frequency. Hence, Figure 9.33 gives a graphical explanation. If

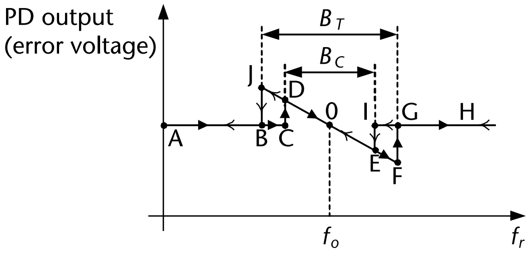


**Figure 9.33** Phase error signal as function of the reference signal frequency.

the frequency of the reference signal varies from zero to infinite, the operation point of the error signal runs across path ABCD0EFGH and vice versa if the frequency decreases. As long as it runs on the segment JDEF, the PLL is locked.

### 9.7.3 PLL Linear Analysis, Stability

PLLs are inherently nonlinear. The main source of nonlinearity is the PD. However, using a small signal assumption and the fact that the multiplication operation in PD is continuous, the basic PLL can be approximated—when the loop is locked with a small tracking error—by a second-order linear system similar to a position servo.

If the reference signal is a sine wave (Figure 9.32):

$$r(t) = r_m \sin\left(\omega_r t + \varphi_r\right) \tag{9.75}$$

and the VCO output signal is assumed to be:

$$o(t) = o_m \cos\left(\omega_o t + \varphi_o\right) \tag{9.76}$$

the PD output is the result of a multiplication:

$$e(t) = K \sin\left(\omega_r t + \varphi_r\right)\cos\left(\omega_o t + \varphi_o\right) \tag{9.77}$$

This can be expanded to give:

$$e(t) = \frac{K}{2}\{\sin[(\omega_r + \varphi_o)t + \varphi_r + \varphi_o] + \sin[(\omega_r - \varphi_o)t + \varphi_r - \varphi_o]\} \tag{9.78}$$

The first term on the right-hand side of the expression (9.78) is a high-frequency term that is filtered out by the LPF. The above-mentioned condition for tracking error near to zero allows the assuming: $\omega_o = \omega_r$, and thus the PD output can be approximated by:

$$e(t) = K_{PD}\left(\varphi_r - \varphi_o\right) \tag{9.79}$$

Typically, the LF is a first-order LPF with the transfer function:

$$LF(s) = \frac{K_{LF}}{\tau_{LF} \cdot s + 1} \tag{9.80}$$

where $s$ is the Laplace operator (see (8.49)) and $\tau_{LF}$ is the LF time constant.

Since the phase is the integral of angular velocity, the VCO is modeled as:

$$VCO(s) = \frac{K_O}{s} \tag{9.81}$$

The transfer function of the open-loop gain results combining (9.79) with (9.81):

$$L(s) = \frac{K_L}{s \cdot \left(\tau_{LF} \cdot s + 1\right)} \tag{9.82}$$

where the total gain is:

$$K_L = K_{PD} \cdot K_{LF} \cdot K_O \tag{9.83}$$

which will be recognized as the transfer function of a position servo.

Consequently, the PLL closed-loop transfer function is:

$$T(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \tag{9.84}$$

where the natural frequency $\omega_n$ and the damping $\zeta$ are:

$$\omega_n = \sqrt{\frac{K_L}{\tau_{LF}}} \text{ and } \zeta = \frac{1}{2\sqrt{\tau_{LF} \cdot K_L}} \tag{9.85}$$

In conclusion, the PLL has a standard second-order transfer function and its dynamic behavior and stability are defined by the poles of this function (see Sections 9.1.2.1 and 9.1.2.2).

The block diagram of the linear PLL approximation is shown in Figure 9.34(a). Here $\varphi_r$ and $\varphi_o$ are the reference signal and VCO phase shift, respectively.

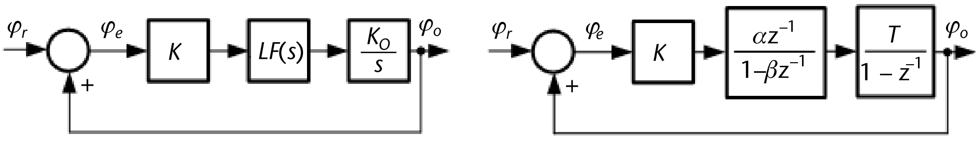
For APPLs, the $B_C$ measures below 1% related to the VCO frequency.

Certainly, a PLL can be implemented using simulation software (e.g., CE2000, ADIsimPLL, etc.).To do this, it has to be put into digital form as shown in Figure 9.34(b). Here, each block in Figure 9.34(a) has to be replaced by its discrete equivalent.

### 9.7.4 Digital PLLs

By inserting an analog-to-digital convertor (ADC), an APLL can be moved into the discrete-time domain, which leads to the digital PLL configuration (Figure 9.35). The digital PLL (DPLL) was developed as an extension of the APLL and is therefore similar in structure to the APLL (Figure 9.32). In this diagram, $n$ represents the discrete-time equivalent of continuous time $t$.

Note that there are strong similarities between Figure 9.32 and Figure 9.35. The DPLL can be extended to measure the phase of a signal consisting not just of fundamental component, but also of harmonically related components.



(a) (b)

**Figure 9.34** Linear approximation of a PLL (a) and digital simulation model (b).

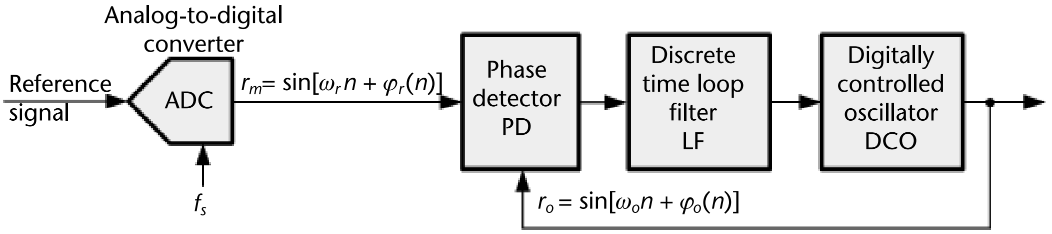

Analog-to-digital converter
Reference signal
ADC
$f_s$
$r_m = \sin[\omega_r n + \varphi_r(n)]$
Phase detector PD
Discrete time loop filter LF
Digitally controlled oscillator DCO
$r_o = \sin[\omega_o n + \varphi_o(n)]$

**Figure 9.35** Basic DPLL for tracking the phase of a discrete-time carrier signal.

The DPLL will function effectively whether the phase is constant or time-varying. The DPLL will introduce minimal processing error (accuracies approaching 0.001°). The main error arises from the ADC inaccuracy.

The PLL is one of the most versatile circuits in electronics today. While it was not widely used for many years, the advent of the IC meant that PLL and synthesizer chips became widely available. This made them cheap to use and their advantages could be exploited to the fullest.

Making exception to the integrated PLLs for very high working frequencies (e.g., ADF4106 and 4153), which are overdesigned for inductive sensor applications, a suitable PLL for this use remains the traditional part CD4046 (manufacturers Texas Instruments, Nexperia) [24, 25], which is a CMOS micropower PLL with a VCO for $f < 2$ MHz, two optional PDs, and an external low-pass RC filter.

For higher operating frequencies up to 17 MHz (sufficient for inductive sensors), a good choice is 74HCT9046A (manufacturer Nexperia) [26]. It is suitable for low supply voltage of 5V and has a similar architecture (Figure 9.36).

Both devices have the possibility to realize the aforementioned very important frequency synthesis. Thus, a digital divider circuit will be inserted between the VCO output and the PD input (pins 4 and 3 in Figure 9.36). This additional hardware divides the VCO frequency by an adjustable factor $N$. By applying an input frequency $f_r$ (pin 14), the PLL makes sure that the divided VCO frequency equals $f_r$, which makes sure that the VCO generates a stable frequency at multiples of an input frequency:

$$f_o = N \cdot f_r \tag{9.86}$$

When the number $N$ is constant, the circuit is called integer-N-PLL; if the circuit can also support variable numbers, it is called fractional-N-PLL.

The PLL's engineering is very comprehensive and this book is limited in size. Hence, for a thorough knowledge of PLLs, the author suggests two basic references: [10] and [22].

*Example 9.4*

The knowledge gained in Sections 9.2, 9.3, and 9.7 can now be used to enumerate three accuracy-levels for the detection of the LC oscillator output signals in inductive sensors resulting from the practical author's experience with detection methods:

1. *Traditional concept.* The popular solution consists of a peak detector (see Section 9.2.2), which is connected to the output LC of an inductive sensor

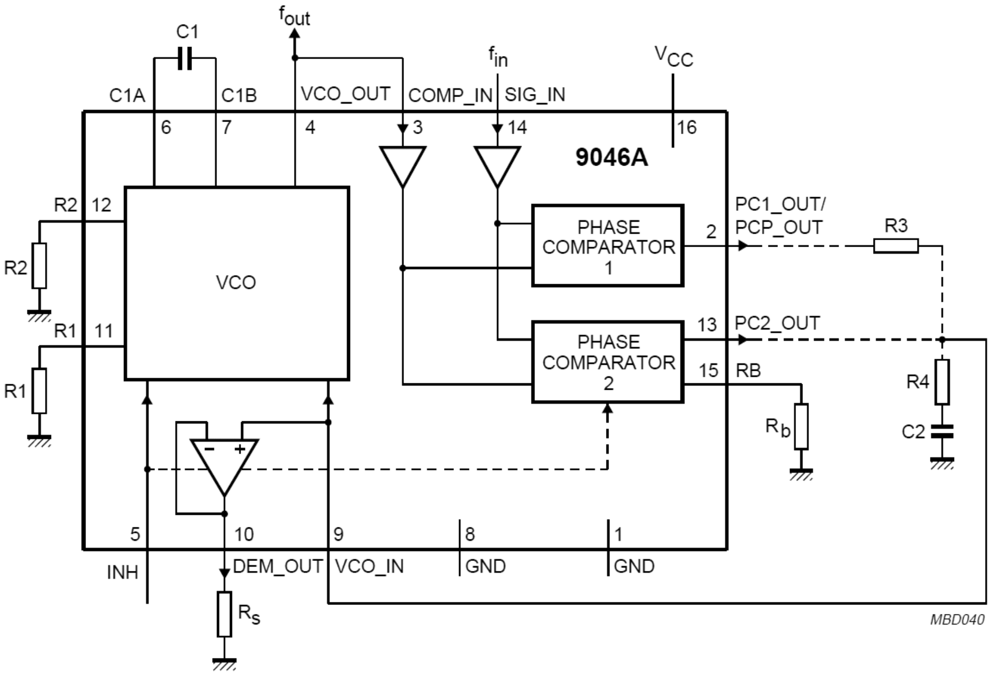


**Figure 9.36** Block diagram of the PLL-IC: CD9046A [26].

oscillator (e.g., in Figure 8.34(b)). The link is direct or through a voltage follower to prevent the oscillator from overloading (Figure 9.37(a)). The solution requires a minimal number of components and therefore is cost-efficient and space-saving. Such applications and their restrictions were presented in previous sections (Figures 7.3, 7.15) and do not need additional information.

2. *Intermediate level.* The significant change is the replacement of the peak detector by a more accurate detector, namely the full-wave rectifier (see Section 9.2.1.2) with a filtering capacitor $C_R$ (Figure 9.37(b)). It is a circuit that provides the absolute value $|v_{LC}|$ of the oscillator output voltage with a lower ripple and a better dynamic. The schematic looks similar to the previous, however, it requires higher costs and more space in the sensor housing.
3. *High level.* This is a mixed-mode detection that combines the accuracy of the analog sample-and-hold procedure with the benefits of a digital sampling with multiplied frequency (Figure 9.37(c)). The sinusoidal oscillator signal $v_{LC}$ is led to the S/H stage (see Section 9.3) and the strobe signal should mark the positive peak of $v_{LC}$ as precisely as possible. Generally, a zero crossing detection offers a high precision.

Zero crossing detections usually offer the best precision. For this reason, the digital path receives the same oscillator signal $v_{LC}$ and realizes a frequency multiplication with a PLL ($N = 2$). The following pulse former circuit generates short strobe pulses with a double frequency by detecting the negative zero crossings of its input and additionally eliminates the odd pulses in pulse sequence. Hence, the strobe signal for S/H is a rectangular wave having only pulses that are synchronous

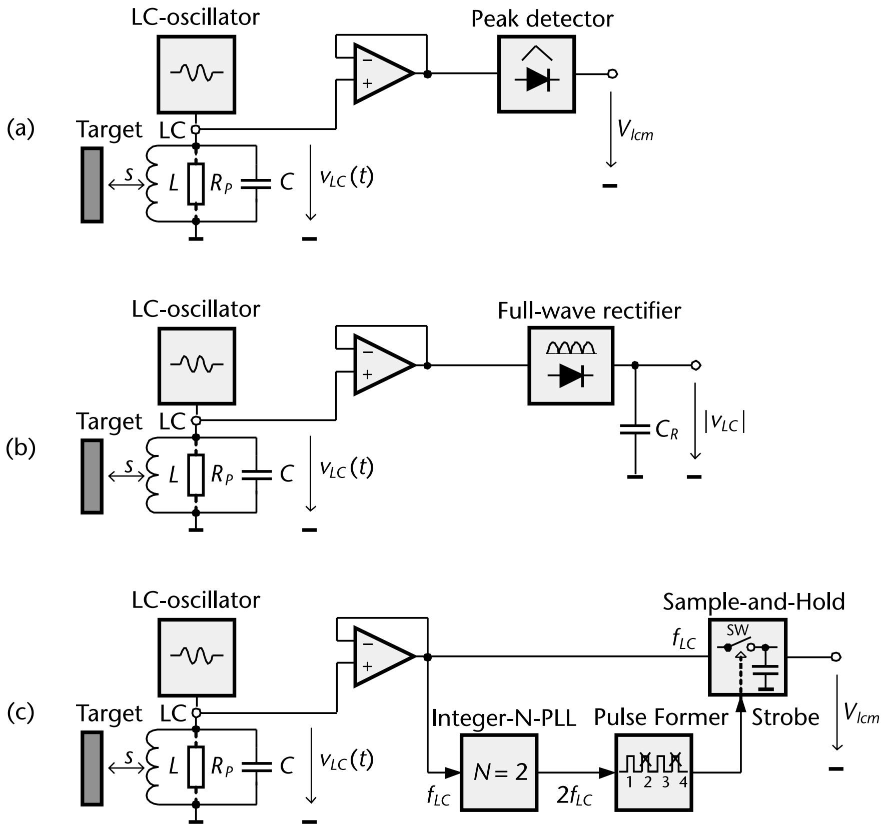

(a)
LC-oscillator
Peak detector
Target LC
s
L
$R_P$
C
$v_{LC}(t)$
$V_{lcm}$
(b)
LC-oscillator
Full-wave rectifier
Target LC
s
L
$R_P$
C
$v_{LC}(t)$
$C_R$
$|v_{LC}|$
(c)
LC-oscillator
Sample-and-Hold
$f_{LC}$
SW
Target LC
s
L
$R_P$
C
$v_{LC}(t)$
Integer-N-PLL
Pulse Former
Strobe
$V_{lcm}$
N = 2
$f_{LC}$
$2f_{LC}$
1 2 3 4

**Figure 9.37** Summary of three established levels for the signal detection in inductive sensors with (a) peak detector, (b) precision rectifier, and (c) sample-and-hold.

with the positive amplitude pecks of the oscillator signal $v_{LC}$. This accurate circuit implies higher hardware investments.

## 9.8 Digital-to-Analog and Analog-to-Digital Convertors

The miniaturization and the cost optimization of the electronic components is a key for the replacement in the sensor electronics of analog signal processing methods by digital methods. Obviously, digital procedures provide significantly higher accuracy, better features, and also additional functions for sensor parameterizing, calibration, temperature compensation, diagnosis, and so forth.

But the nature and all quantities and effects to be detected and measured were, are, and will be analog. For this reason, devices to convert analog values into digital signals or to reconvert digital quantities into analog magnitudes are sine qua non circuits for digitalized sensors.

### 9.8.1 Digital-to-Analog Converters

DACs accept a digital word as an input and convert it to an analog voltage or current. The digital word is presented in different codes, the most common being pure binary or binary-coded-decimal (BCD). For a pure binary codification of the digital input, the analog output can be theoretically expressed by:

$$V_{\text{OUT}} = \sum_{n=0}^{N-1} \left(2^n \cdot a_n\right) \cdot V \tag{9.87}$$

where the voltage $V$ is a proportionality factor and the coefficients $a_n$, $a_{n-1}$, ... $a_0$ represent the binary word and $a_n = 1(0)$ if the $n$th bit is 1(0). The most-significant bit (MSB) is that corresponding to $a_{N-1}$ and the least-significant bit (LSB) corresponds to $a_0$.

The implementation of (9.87) leads to one of the most popular DACs, namely the binary-weighted DAC (Figure 9.38). It consists of an inverting op-amp working as current-to-voltage convertor, which has a fixed feedback resistor $R_R$ and a resistor network R-2R-4R...$2^{N-1}$R in the input path. The parallel resistors are connected to GND or to a reference voltage $V_{REF}$ (usually of −10V) by means of single-pole double-throw electronic switches depending on whether the corresponding bit is 0 or 1.

Thus, for a number $N$ of bits, which are binary-weighted as in (9.87) the real DAC voltage results in:

$$V_{\text{OUT}} = \frac{V_{REF}}{2^{N-1}} \cdot \frac{R_R}{R} \cdot \sum_{n=0}^{N-1} \left(2^n \cdot a_n\right) \tag{9.88}$$

Many implementations for the digitally controlled switches are possible, such as totem-pole MOSFET drivers or CMOS inverter feeding an op-amp.

The DAC requires a low number of resistors, but has limited accuracy and stability, which depend on the quality and temperature stability of the resistors.

For a high-digit number it is difficult to cover the resistance range. If the base value $R$ is low enough (to obtain a realistic highest $2^{N-1}R$ value), the $R_{\text{ON}}$ value of the switches becomes comparable and affects the conversion accuracy and linearity [6]. In contrast, for higher $R$ base value, the resulting large $2^{N-1}R$ resistors cannot be

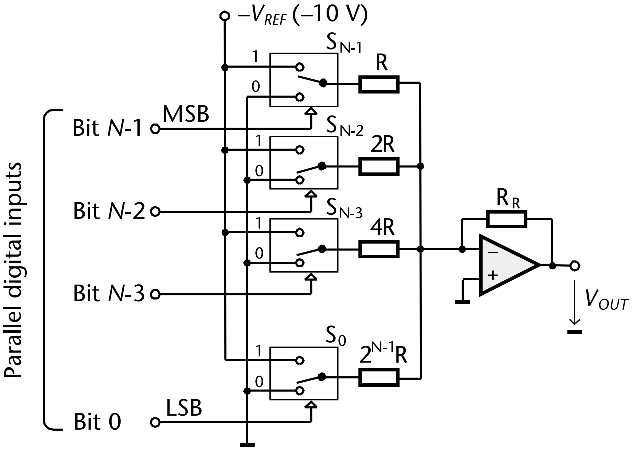


**Figure 9.38** A DAC with binary-weighted resistors.

economically integrated in silicon. Consequently, this type of ADC is preferably used for application-specific circuits and seldom were more than 4 to 6 bits are required.

The difficulties of extreme resistance values are avoided by the ladder-type DAC, which is frequently used in high-resolution data-conversion systems. The circuit utilizes twice the number of resistors for the same number $N$ of bits but of values $R$ and $2R$ only. The DAC consists of a noninverting op-amp with fixed amplification and the resistor network (ladder) (Figure 9.39).

The ladder used is a current-splitting device, and thus the ratio of the resistors is more important than their value. This can be better controlled by matching techniques in the design and process of integrated circuits. At any of the ladder nodes the resistance is $2R$ looking to the left or to the right or toward the switch. As a result of this property, if just the MSB switch $S_{N-1}$ is closed, the resulting voltage at the noninverting op-amp terminal is $V_{REF}/3$ and the output voltage is:

$$V_{\mathrm{OUT}} = \frac{V_{REF}}{3} \cdot \frac{R_1 + R_2}{R_1} \tag{9.89}$$

Furthermore, applying the principle of superposition, the output voltage can be generally expressed by:

$$V_{\mathrm{OUT}} = \frac{V_{REF}}{3 \cdot 2^{N-1}} \cdot \frac{R_1 + R_2}{R_1} \cdot \sum_{n=0}^{N-1} \left(2^n \cdot a_n\right) \tag{9.90}$$

Because of the stray capacitance from the nodes to ground, there is a propagation time delay [6], from left to right down the ladder. When the LSB switch closes, this time delay is much longer than the MSB switch closes. Hence, when the digital word changes, a transient waveform will appear at the output before $V_{\mathrm{OUT}}$ settles down to its proper value.

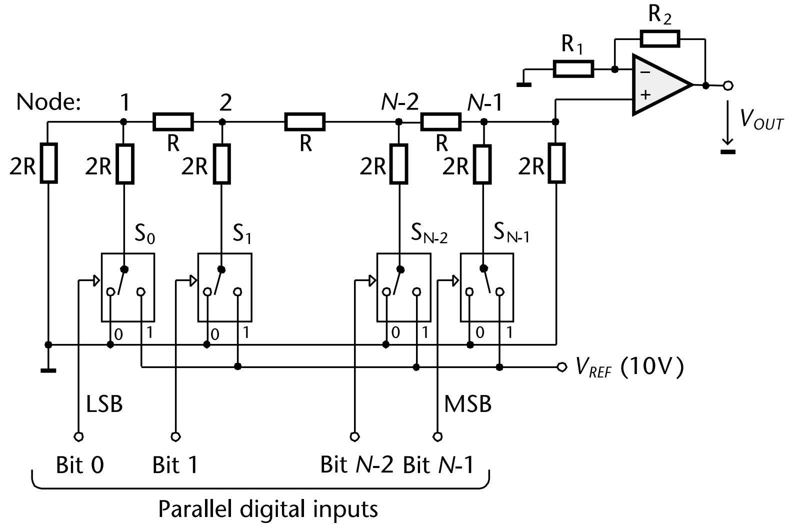


**Figure 9.39** An R–2R ladder DAC.

The basic DACs described in this section must be augmented by additional circuitry as latches and logic circuits are needed to input the digital data. Monolithic systems containing all these schematics are available.

For inductive sensor applications, DACs with 8-bit or 12-bit resolution should be sufficiently precise. Such convertors are manufactured by several companies. For example, Analog Devices offers AD558 and AD7541, which are 8- or 12-bit DACs with *R*-2*R* ladder. A good information source for the availability of DACs (8 to 20 bits) from Analog Devices, Maxim, Burr-Brown, Plessey, and others, is the Artech volume [6].

### 9.8.2 Analog-to-Digital Converters

The usual method to bring analog signals into a sensor digital signal processing unit is to use an ADC. This receives an input voltage or a current and converts it to a digital output.

The ADC generally accepts two voltages: the effective input voltage $V_{\mathrm{IN}}$ and a reference voltage $V_{REF}$, representing the maximum value that the ADC can convert. For a specified ADC bit number $N$, $V_{REF}$ determines the step-size or converter's resolution, which is the smallest voltage change that can be measured by the ADC:

$$STS_{ADC} = \frac{V_{REF}}{2^N} \tag{9.91}$$

The resolution can be improved by reducing $V_{REF}$, which actually means an unwanted limitation of the measuring range. The suitable way to increase the resolution remains to grow the number $N$.

ADCs come in various speeds and degrees of accuracy. The most popular types are

- Tracking ADC;
- Successive approximation ADC;
- Flash ADC;
- Dual-slope ADC (ratiometric converter);
- Sigma-delta ADC.

Tracking ADCs have an analog comparator, an up/down counter, and a DAC, together forming a closed-loop system (Figure 9.40(a)). The comparator compares the input voltage with the DAC output voltage. If the input is higher/lower than this voltage, the counter counts up or down. The DAC, which is connected to the counter output, continuously tracks the counter and brings its content in analog form at the comparator input.

The primary drawback of the tracking procedure is the speed. The maximum clock speed depends on the DAC propagation delay and comparator response time. The method can be implemented with discrete hardware or in specific IC rather than in commercial parts.

Successive approximation ADCs are similar to tracking ADCs, execpt that their successive approximation register SAR performs a binary search instead of

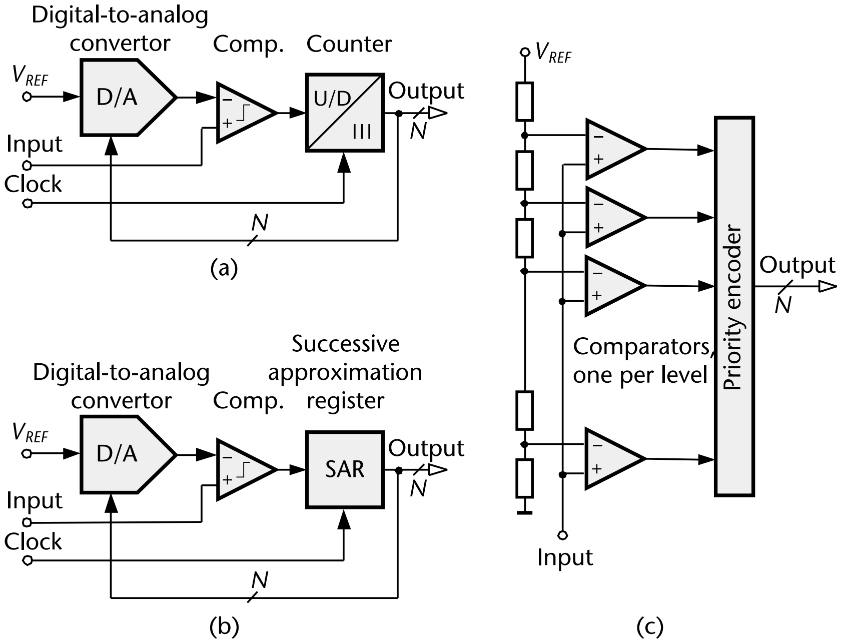

Digital-to-analog convertor
Comp.
Counter
$V_{REF}$
D/A
U/D
Output
N
Input
Clock
N
(a)
Successive approximation register
Digital-to-analog convertor
Comp.
$V_{REF}$
D/A
SAR
Output
N
Input
Clock
N
(b)
$V_{REF}$
Priority encoder
Output
N
Comparators, one per level
Input
(c)

**Figure 9.40** Block diagrams of (a) tracking ADC, (b) successive approximation ADC, and (c) flash ADC.

just counting up or down by 1 (Figure 9.40(b)). An $N$-bit successive approximation ADC needs only $N$ clocks to convert the input voltage. The operation flow is a string of the following steps:

- Set successively the SAR bits (starting with MSB) and look at the comparator output;
- If this output is high (DAC output is lower than the input voltage), leave the corresponding bit set and continue with the next inferior bit;
- If this output goes low (DAC output becomes higher than the input), reset this bit and continue with the next inferior bit;
- At the end, a DAC output voltage results, which is as closed as possible to the input voltage. The final digital word is available at the SAR output.

These ADCs require more logic, but the conversion is consistent and usually faster.

Flash ADCs are the fastest type available. They have one comparator per voltage step. Hence, an $N$-bit flash ADC has $2^N$ comparators (Figure 9.40(c)). The noninverting inputs of all comparators are connected to the ADC input. The inverting input of each comparator is connected to one point in a string of resistors (voltage divider). The comparator outputs are connected to a block of logic ($2^N$ inputs and $N$ outputs) that determines the digital word at its output based on which comparators are low or high.

The high conversion speed is the sum of one comparator delay and the logic delay (usually negligible). Flash ADCs are very fast, but due to the high number of comparators, they require a large silicon area and have a high current consumption.

Dual-Slope ADCs use a series chain consisting of an integrator (see Figure 9.2(b)), a comparator, a gate for the clock pulses, and a counting logic (Figure 9.41(a)).

The integrator input is first switched to the input signal $v_{\mathrm{IN}}$ and the integrator output charges toward this input voltage. In this first conversion phase, the gate is open and the counter counts up the clock pulses.

After a specific number, a clock cycles $n_{\mathrm{MAX}}$ (integration time), the counter becomes full, the control logic switches the integrator input from $v_{\mathrm{IN}}$ to a reference voltage $V_{REF}$, and the integrator charges down to zero. Its output remains HIGH, the gate is further opened, and the already reset counter starts using the same clock.

During this second conversion phase, when the integrator output obtains zero volts, the comparator output goes LOW, the counter stops, and the new clock pulses number $n_{\mathrm{IN}}$ is recorded.

The control unit computes the input voltage using the conversion equation of this ADC-type [6]:

$$V_{\mathrm{IN}} = \frac{n_{\mathrm{IN}}}{n_{\mathrm{MAX}}} \cdot V_{REF} \tag{9.92}$$

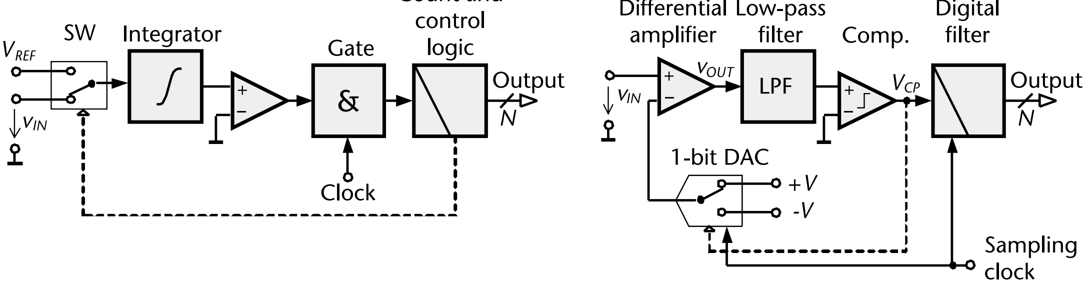
and converts this information into the digital word.

A single-slope conversion is possible. It runs the counter while charging up and stops counting when a reference voltage is reached (instead of charging for a specific time). However, this procedure is affected by clock accuracy. The dual-slope conversion eliminates clock accuracy problems because the same clock is used to count both $n_{\mathrm{MAX}}$ and $n_{\mathrm{IN}}$.

The dual-slope ADC takes a relatively long time to perform the conversion, but the inherent filtering action of the integrator eliminates noise. Clock jitters does not affect accuracy.

Sigma-delta ADCs have an input signal path consisting of a differential amplifier, an LPF (integrator, third to fifth order), and a comparator with grounded inverting input (Figure 9.41.(b)) [27].

Hence, if the LPF output is above ground, the comparator output $V_{CP}$ will be HIGH and vice versa. This output drives a digital filter and a 1-bit DAC, whose



(a) (b)

**Figure 9.41** Block diagrams of a (a) dual-slope ADC and (b) sigma-delta ADC.

output switches between two voltages: $+V$ (for DAC control input on HIGH) and $-V$, which are + full-scale and – full-scale.

Finally, the DAC output supplies the inverting input of the differential amplifier and so its output voltage is:

$$v_{\mathrm{OUT}} = \begin{cases} v_{\mathrm{IN}} + V & \text{for } V_{CP} \rightarrow \mathrm{HIGH} \\ v_{\mathrm{IN}} - V & \text{for } V_{CP} \rightarrow \mathrm{LOW} \end{cases} \tag{9.93}$$

Consequently, the duty cycle of the DAC output represents the input level. The digital filter converts this modulated signal to a binary value.

The input range of the sigma-delta DAC is the plus-and-minus DAC voltage.

The primary advantages of the sigma-delta DAC are the high resolution and theoretically perfect linearity [27]. The system is suitable for oversampling; that is, it can sample the signal several times faster than data are actually needed. In this way, by averaging samples the accuracy increases remarkably.

In opposition to the previously described convertors, the sigma-delta DAC does not have any resistors, which could affect the accuracy due to mismatch or temperature drift.

The disadvantage is the speed. Because the converter works by oversampling the input, the conversion takes many clocks and makes the convertor slower than other types. Another drawback is the complexity of the digital filter that converts the duty cycle information to a digital word.

Making a benchmarking of the above-described ADCs, the sigma-delta DAC holds a top position. Regarding the conversion speed, it reaches the range of successive approximation ADCs (≤ 100 kHz), but is not as fast as even the slowest flash ADC. With respect to the resolution, the sigma-delta DAC is the winner, offering a resolution up to 24 bits.

From an inductive sensor-specific viewpoint, successive approximation ADCs (e.g., ADS7805, AD7621) provide good compromise results: middle resolution (8 to 16 bits) and high speed (50 kHz to 10 MHz) and hence could be considered as a favorite solution for these sensors.

The majority of IC manufacturers possess ADC-design macros that are integrated in commercial devices and systems on chip but can be also integrated in customer ASICs. Dual-slope and sigma-delta are the preferred offers [28].

Finally, for a deeper insight into the DAC and ADC topic, the author invites the reader to study the Artech reference [6]. Good summaries of the available commercial convertors are available here.

## References

[1] Jung, W. (ed.), *Op-Amp Application Handbook*, Amsterdam: Newnes/Elsevier, 2006.

[2] Amos, S., and M. James, *Principles of Transistor Circuits*, ebook, Ninth Edition, Oxford, UK: Newnes, 2003.

[3] Christiansen, D., and A. Charles, *Standard Handbook of Electronic Engineering, Fifth Edition*, New York: McGraw-Hill Professional, 2005.

[4] Mancini R., *Op-Amps for Everyone*, web.mit.edu/6.101/www/reference/op_amps_everyone.pdf, Texas Instruments, 2002.

[5] Wilson, J., et al., *Test and Measurement–Know It All*, Burlington, MA: Elsevier / Newnes, 2017.

[6] Nawrocki, W., *Measurement Systems and Sensors*, Second Edition, Norwood, MA: Artech House, 2016.

[7] Camenzind, H., *Designing Analog Chips*, ebook, Book Surge Publishing, 2005.

[8] www.analog.com/media/en/technical-documentation/.../OP07.pdf, Datasheet OP07—Analog Devices, Rev. C, 2011.

[9] Becker, W. J., K. W. Bonfig, and K. Hoeing (eds.), *Handbuch elektrische Messtechnik*, 2nd Edition, Heidelberg, Germany: Huethig-Verlag, 2000.

[10] Tietze, U., C. Schenk, and E. Gamm, *Halbleiter-Schaltungstechnik*, 15th Edition, Berlin, Germany: Spinger/Vieweg-Verlag, 2016.

[11] Rahal, M., and A. Demosthenous, "A Synchronous Chopping Demodulator and Implementation for High-Frequency Inductive Position Sensors," *IEEE Transactions on Instrumentation and Measurement*, Vol. 58, No. 10, October 2009.

[12] Loos, H. R., *Systemtechnik induktiver Weg- und Kraftaufnehmer*, Ehningen, Germany: Expert Verlag, 1992.

[13] http://www.analog.com/media/en/training-seminars/tutorials/MT-079.pdf, Tutorial: MT-079: Analog Multipliers, Rev.0, October 2008.

[14] Gray, P. R., P. J. Hurst, H. S. Lewis, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, Fifth Edition, New York: John Wiley & Sons, 2009.

[15] http://www.analog.com/media/en/training-seminars/tutorials/MT-079.pdf, Tutorial–Applications: Analog Multipliers.

[16] www.ti.com/lit/ds/symlink/lf398-n.pdf, Datasheet: LF198/LF298/LF398, Monolithic Sample-and-Hold Circuits, 2015.

[17] Pasic, A., and J. Dowling, "Linearizing Calibration Methods for a Generic Embedded Sensor Interface (GESI)," 1st International Conference of Sensing Technology, Palmerston, New Zealand, 2005.

[18] Vemuri, A., and C. Tong, "Linearization of Sense Element Outputs Using Sensor Signal Conditioners," *EDN Network*, 2016.

[19] Sturcel, J., and Kamensky, M., "Function Approximation and Digital Linearization in Sensor Systems," *AT & P Journal PLUS2*, 2006.

[20] Medrano-Marques, N. J., and B. Martin-del-Brio, *A General Method for Sensor Linearization Based on Neural Networks*, ISCAS 2000, Geneva, Switzerland, 2000.

[21] http://www.ti.com/lit/ds/symlink/lm211.pdf, Datasheet Differential Comparators LM111, 211 and 311, March 2017.

[22] Best, R. E., *Phase-Locked Loops, Theory Design, and Applications*, 6th Edition, New York: McGraw-Hill, 2007.

[23] Kruse, J., Tutorial: Phase Locked Loop Seminarvortrag, https://documentde.com/phase-locked-loop-phasenregelschleife.

[24] www.ti.com/lit/an/scha002a/scha002a.pdf, David Morgan, CD4046B PLL Application Report, Texas Instruments, 2003.

[25] https://assets.nexperia.com/documents/data-sheet/HEF4046B.pdf, HEF4046B Phase-locked loop, Nexperia, 2016.

[26] https://assets.nexperia.com/documents/data.../74HCT9046A.pdf, 74HCT9046A PLL with band gap controlled VCO, Nexperia, 2016.

[27] Webster, J. G. (ed.), *The Measurement, Instrumentation and Sensors Handbook*, Boca Raton, FL: CRC Press/IEEE Press, 1999.

[28] http://www.ti.com/data-converters/adc-circuit/overview.html, Overview Data-Converters, Texas Instruments, 2017.

## CHAPTER 10

# Inductive Sensors: Output Signal Providing

Chapter 10 completes the block-by-block presentation of the signal chain in inductive sensors. This presentation started with a functional block diagram (Figure 7.1) showing the generic structure of the evaluation electronics in inductive sensors. The basics, operation, and realizations of the stages in the signal chain were subjects of the previous chapters.

We achieve now the final positions of the BEE, marked by the analog output stage (AOS) for output signals of analog sensors and switching output driver (SOD) for binary output signals.

These analog and binary signals are already defined and described (physical scale, classification, values, etc.) in Sections 1.3.7.1 and 1.3.7.2, respectively. If necessary, the reader is invited to review this knowledge. Digital forms of the sensor outputs defined in Sections 1.3.7.4 and 1.3.7.5 will be treated in Chapter 14.

Chapter 10 deals with methods and circuits to generate analog binary output signals and examines conventional commercial output drivers for analog and digital inductive sensors. Several generations with their features of output stages for inductive sensors are presented.

Some integrated output stages were described in the previous chapters (e.g., Figures 7.21 and 7.22).

## 10.1 Output Stages for Analog Inductive Sensors

To reiterate knowledge regarding the sensor output types (Section 1.3.7), the electrical data domain of sensor outputs—the name of a quantity used to represent or transmit information—covers three subdomains, the first being the analog domain, in which the information is carried by signal amplitude (usually voltage or current). The analog domain is the most popular, versatile, and service-friendly electrical domain for sensor outputs, even though some people consider it to be technologically outdated.

### 10.1.1 Voltage Output Stages

The voltage as physical support for sensor output signals is used in a wide range of applications. That is the reality in electronic engineering: voltage use forwards current use.

In principle, a good choice for a voltage output stage is the op-amp (Section 9.1), as long as its output has sufficient current capability.

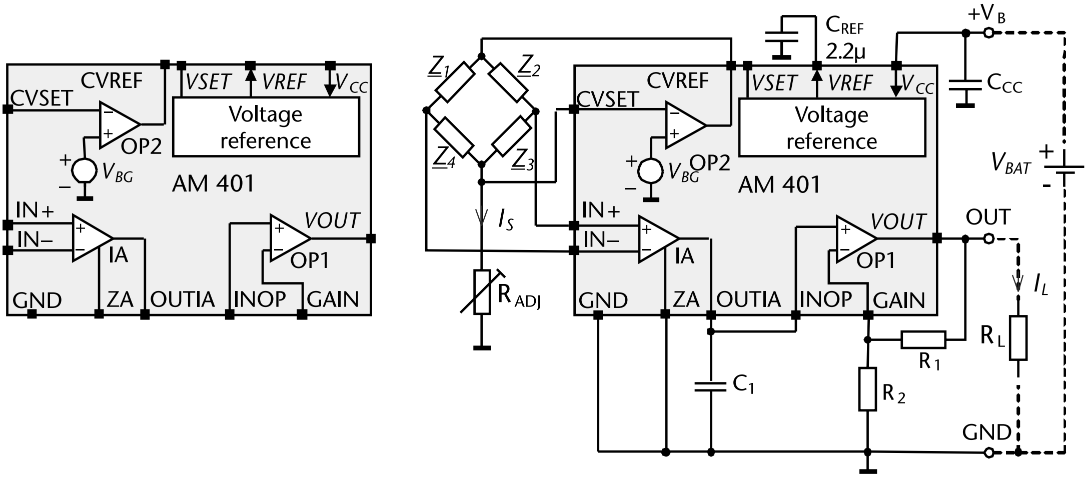
According to the standards, the load resistance of voltage output stages has a minimum value of $R_{Lmin} = 1\ \mathrm{k\Omega}$, and hence, the maximum output current for standard output voltage range $V_{\mathrm{OUT}} = 0$ to 10V achieves values, which are today inside the specification of the majority of op-amps:

$$I_{\mathrm{OUT}} \leq \frac{V_{\mathrm{OUT}}}{R_{Lmin}} = 10\ \mathrm{mA} \tag{10.1}$$

The drawback of this simple low-cost version is the missing protection of its output against overload or short circuit, overvoltages resulting by switching off of inductive loads, EMC parasitic signals, and so forth.

These disadvantages are eliminated if an integrated specific device is applied. The commercial part AM401 [1] is a representative output stage with various application possibilities for industrial sensors. The modular monolithic device consists of the following four parts with individual configuration options (Figure 10.1(a)):

- A robust voltage regulator with selectable regulated voltage $V_{REF}$ = 10V or 5V;
- An op-amp (OP1) for input signals referenced to GND, suitable for analog output stages and having an adjustable gain;
- A high-precision instrumentation amplifier (IA) with an internal gain of $G_{IA} = 5$ and the ability to externally set a reference potential (input ZA);
- A second noninverting op-amp (OP2), which is biased by a bandgap voltage reference $V_{BG}$ = 1.27V, typ. [2], and can be used as a current or voltage source to supply external additional circuits.



(a) (b)

**Figure 10.1** Integrated voltage output stage AM401: (a) functional block diagram, and (b) application with current supplying of an inductive measuring bridge.

Additional technical data includes

- Supply voltage range: $V_{CC} = 6$ to 35V, quiescent current ≤ 1.5 mA;
- Regulator current capability: $I_{REF} \leq 10$ mA;
- IA-differential input voltage range: $V_{\mathrm{IN}+} - V_{\mathrm{IN}-} = \pm 400$ mV;
- OP1-input voltage range: $V_{INOP} = 0$ to 5 V;
- OP1-output voltage range: $V_{\mathrm{OUT}} = 0$ to $V_{CC} - 5$V;
- Load resistance: $R_{L\_\mathrm{OUT}} \geq 2$ kΩ;
- Operating temperature range: −40°C to +85°C;
- Protection against reverse polarity;
- Output current limitation (10 mA), short-circuit protection;
- Protection against ESD.

The AM401 architecture, operation, and basic application for 0 to 10V output was already presented in Section 7.4.2 (Figure 7.22). In this case, the voltage reference of AM401 supplies the front-end of an inductor sensor. The sensor output is processed by OP1 and converted to the standard output voltage. The stages IA and OP2 were not used and correspondingly disabled.

To illustrate the versatility of AM401, Figure 10.1(b) shows a different application consisting of the current supplying of a deflection bridge and amplification of the signal in its diagonal branch. Thus, the OP2 is connected as current source, which is referenced by $V_{BG}$ and delivers a supply current $I_S$ for the measurement bridge. The current is set by $R_{ADJ}$:

$$I_S = \frac{V_{BG}}{R_{ADJ}} \tag{10.2}$$

The typical applications list also contains an analog output stage for DACs and MCUs with internal DAC, impedance convertor, and so forth.

### 10.1.2 Current Telemetry: Current Output Stages

In current telemetry the sensor signal is converted to a proportional current, which is sent to the connecting lines. The receiver detects this current by measuring the drop in voltage across a known resistor. Usual current values are 0 to 20 mA and 4 to 20 mA (by far the most common) but also 0 to 5 mA, 1 to 5 mA, and so forth. [3].

Although voltage signals are very widely used, the signal transmission over longer distances and electromagnetic-disturbed environments is more reliable using the current carrier. The signal is not altered by the voltage drop over the wires, contact resistances, or temperature drift. The telemetry is insensitive to parasitic thermocouples. The internal current source has a high impedance, thus induced interference voltages cannot generate parasitic currents. As well, it is possible to connect more receiving users in series along the current loop.

Most used are systems with the 4- to 20-mA current loops. Beside the abovementioned advantages, the loop 4 to 20 mA has specific beneficial properties because it allows to distinguish between zero signal (4 mA) and no signal (0 mA) as a result of an unwanted cable broken (open-wire failure). If it is necessary to reduce the wire numbers, the zero current can be used to supply the sensor.

Investigating conventional schematics, we find two solutions to achieve the range with 4-mA offset: (1) to add an offset voltage to the input voltage and to convert the sum by a voltage-controlled current source, or (2) to primarily convert the input voltage into current, to add to this result an offset current, and eventually, to make a final current amplification.

Figure 10.2 shows the schematic of a discrete current output stage according to the first option.

The adder-subtractor with op-amp OA1 (see Figure 9.2(a)) realizes together with the transistor $T_1$ a voltage-controlled current source. An offset voltage of 0.5V is buffered by the follower with OA2 and added to the differential input voltage $V_{\text{IN}}$. Due to the negative feedback, for a deviation $V_{\text{IN}} = 0$ to 10V, the circuit forces a variation 0.5 to 2.5V at the emitter of $T_1$. The collector current of $T_1$ is the loop current and has a value, which is nearly independent of the loop resistance $R_{LOOP}$, as long as this is a low resistance. According to the current expression, its following range results in:

$$I_{\text{LOOP}} \approx \frac{V_{RE}}{R_E} = \frac{0.5\text{V to }2.5\text{V}}{125\Omega} = 4\text{ mA to }20\text{ mA} \tag{10.3}$$

The typical system setup is $R_{LOOP} = 250\Omega$ and $V_{CC} = 24\text{V}$. The resulting current should remain unaffected by $R_{LOOP}$ up to 600Ω. The open-collector output $I_{\text{OUT}}$ (Figure 10.2) is very flexible and also supports higher loop resistances, if the supply voltage $V_{CC}$ is correspondingly increased.

However, the use of integrated devices, with their protection functions, is also in this case a preferred option. The equivalent of AM401 (Figure 10.1) for current output is AM462 (Figure 10.3).

The modular monolithic device consists of the following four parts with individual configuration options (Figure 10.3(a)):

- The voltage regulator and the noninverting op-amp OP2, which are identical to those for AM410;

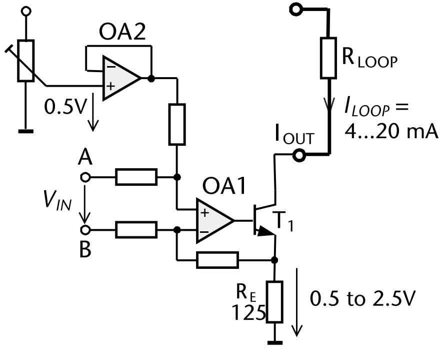
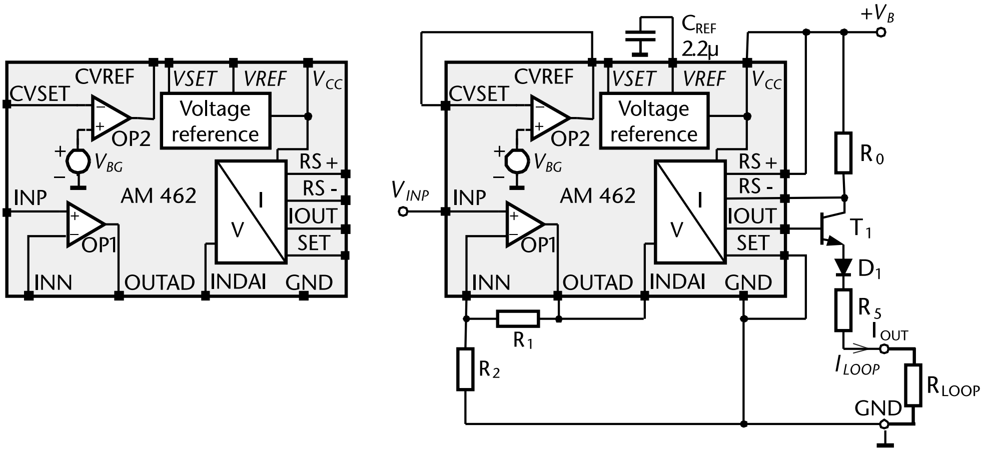


**Figure 10.2** Current output stage implemented with op-amps.



(a) (b)

**Figure 10.3** Integrated current output stage AM462: (a) functional block diagram, and (b) typical three-wire application for 0 to 20 mA.

- An op-amp OP1 (similar to OP1 of AM401), whose gain can be externally set;
- The specific voltage-current convertor that needs an external transistor to perform the V/I conversion.

Additional technical data [4]:

- Supply voltage range: $V_{CC} = 6$ to 35V, quiescent current ≤ 1.5 mA;
- Fixed reference voltages: 5V (pin VSET open) or 10V (VSET to GND) or adjustable between 4.5 and 10V with an external voltage divider;
- Regulator current capability: $I_{REF} \leq 10$ mA;
- Internal bandgap reference: 1.27V, typ.;
- Industrial current output: 0/4 to 20 mA;
- Transconductance of V/I convertor (Table 9.1): $G_m = 0.125$ mA/V;
- Load resistance: $R_{LOOP} = 0$ to 600Ω;
- Operating temperature range: −40°C to +85°C;
- Protection against reverse polarity;
- Output current limitation (10 mA), short-circuit protection;
- Protection against ESD.

Figure 10.3(b) shows a typical application for 0 to 20 mA, where OP1 is connected as a noninverting amplifier that amplifies the input voltage $V_{INP}$ and controls the V/I convertor. OP2 is not used and $V_{REF}$ is set on 5V. To reduce the power dissipation, the final transistor $T_1$ is an external component and protected against reverse polarity by an additional diode $D_1$. If the offset $I_{SET} = 4$ mA is required, the pin SET must be biased on $V_{SET}$ by means of an external voltage divider connected between VREF and GND.

The expression of the output current is:

$$I_{\text{LOOP}} = \frac{V_{INDAI} \cdot 0.125 \text{ mA/V}}{R_0} + \frac{V_{\text{SET}}}{2 \cdot R_0} \qquad (10.4)$$

where $R_0$ is the external resistor that permits the output current to be finally adjusted. Recommended values: $R_0 = 27\Omega$ and $R_5 = 39\Omega$.

The minimal required value of the supply voltage $V_B$ is dependent on the maximal loop resistance, loop current and on the set $V_{REF}$:

$$V_B \geq I_{\text{LOOPmax}} \cdot R_{\text{LOOP}} + V_{\text{CCmin}} = I_{\text{LOOPmax}} \cdot R_{\text{LOOP}} + \left(V_{REF} + 1V\right) \tag{10.5}$$

Accordingly, the maximal resistance $R_{\text{LOOP}} = 600\Omega$ is allowed only for $V_B >$ 18V (at $V_{REF} = 5$V).

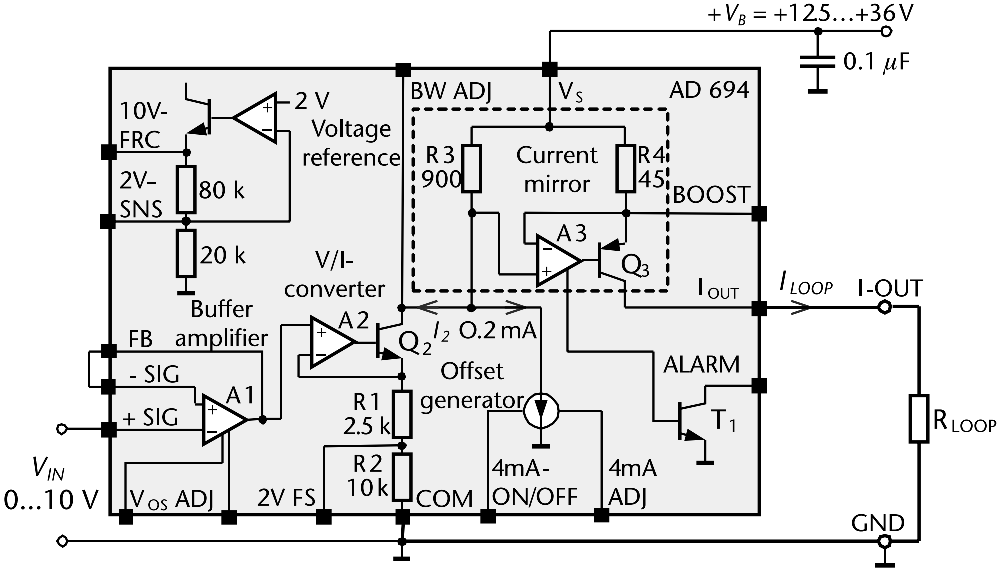
Due to the beneficial modular concept, the design rules regarding the AM401 stages: OP1, OP2, and voltage reference of are valid also for AM462.

The monolithic modular frame concept of Analog Microelectronics [5] offers flexibility of the entire product portfolio. AM401 and AM462 are specialized V or I output stages. Their blocks are combined the other way round to get new applications as

- AM452, which has a AM401-structure (four stages) extended by the V/I convertor of AM462;
- AM400 is the full-power device (six stages), which simultaneously provides both V and I output signals.

The devices of the AM family are available in the shrink small outline package (SSOP16 package) or as dice on 5-inch blue foil, for chip on board assembly with wire bonding (see Section 7.3.1.2).

Obviously, if the power dissipation and thermal conditions inside sensor housing permit, it is more convenient to use circuitry without external transistor (Figure 10.4).



+$V_B$ = +12.5...+36 V
0.1 μF
10V-FRC
2V-SNS
80 k
20 k
2 V
Voltage reference
BW ADJ
Vs
AD 694
R3 900
R4 45
Current mirror
BOOST
A3
Q3
V/I-converter
Buffer amplifier
A2
Q2
I2 0.2 mA
IOUT
ILOOP
I-OUT
FB
- SIG
+ SIG
A1
Offset generator
ALARM
T1
R1 2.5 k
R2 10 k
RLOOP
VIN 0...10 V
VOS ADJ
2V FS
COM
4mA-ON/OFF
4mA ADJ
GND

**Figure 10.4** Integrated current output stage AD694: simplified functional block diagram and typical application according to [6].

Following the signal path in Figure 10.4, the option: offset current addition after voltage conversion can be identified. In fact, the input signal $V_{\mathrm{IN}}$ is buffered by the follower with the amplifier A1 and is converted into an intermediate current by the converter with A2. If the pin 2V FS is open, the collector current of the transistor Q2 is

$$I_2 = \frac{V_{\mathrm{IN}}}{R_1 + R_2} = \frac{V_{\mathrm{IN}}}{12.5\ \mathrm{k\Omega}} \tag{10.6}$$

This current is summed with an offset current of 0.2 mA, and the current sum represents the input current in the current mirror realized with $R_3$ and $R_4$, amplifier A3 and transistor Q3. The current multiplication is given by the ratio $R_3$ to $R_4$, and the output current $I_{LOOP}$ is expressed by:

$$I_{\mathrm{LOOP}} = \left(\frac{V_{\mathrm{IN}}}{R_1 + R_2} + 0.2\ \mathrm{mA}\right)\frac{R_3}{R_4} = 20 \cdot \left(\frac{V_{\mathrm{IN}}}{12.5\ \mathrm{k\Omega}} + 0.2\ \mathrm{mA}\right) \tag{10.7}$$

The device is very versatile and has the following main features [6]:

- Precision voltage reference, programmable to 2.00 or 10.00V;
- Single- or dual-supply operation;
- Power supply range: 4.5 to 36V;
- Open-loop alarm;
- Option for external transistor to reduce self-heating errors;
- Very low nonlinearity 0.002%, typical;
- Available in 16-pin CERDIP and plastic SOIC.

In addition to the described main channel, the device contains a voltage reference, which provides two pin programmable input ranges: 0 to 2V for jumper between pins 10V-FRC (force) and 2V-SNS (sense) or 0 to 10V (open pins). By means of external resistor networks, the reference output can be set to any value between 2 and 10 V [6].

The input amplifier with available feedback output can be used to buffer but also to scale the input voltage.

The transfer function of the output driver can be changed by connecting the pin 2V full scale (FS) to GND. The resulting expression:

$$I_{\mathrm{LOOP}} = \left(\frac{V_{\mathrm{IN}}}{R_1} + 0.2\ \mathrm{mA}\right)\frac{R_3}{R_4} = 20 \cdot \left(\frac{V_{\mathrm{IN}}}{2.5\ \mathrm{k\Omega}} + 0.2\ \mathrm{mA}\right) \tag{10.8}$$

evidences an output current 4 to 20 mA for 0 to 2V input voltage.

The offset generator may be switched off with a TTL control pin (4mA-ON/OFF) to allow 0- to 20-mA operation. The output current offset can be adjusted between 2 and 4.8 mA by a potentiometer connected between the pin 4mA-ADJ and GND [6] and/or with an additional current at the pin BW ADJ (adjust), if any. The bandwidth of the device can be limited to provide noise filtering connecting a capacitor between BW ADJ and $V_S$.

The open collector output ALARM warns of loop failure due to open wires or noncompliance of the output stage. The pin can be tied to an auxiliary supply source (e.g., 5V) through a pullup resistor. The transistor is OFF in normal operation. In the event of a failure, the ALARM pin is driven LOW (allowed alarm current < 20 mA).

The monitored output BOOST is provided to drive an external pass transistor. The feature offloads power dissipation to extend the temperature range of operation and minimize self-heating error. The emitter of an NPN transistor has to be connected to BOOST and its base to the output $I_{\text{OUT}}$. The open collector is the new output, which drives the current in the loop, having the other end biased to a second negative voltage: 0 to –23.5V.

A second similar device AD693 should also be mentioned. It is a loop-powered, 4- to 20-mA sensor transmitter with three stages: (1) an instrumentation amplifier front-end, (2) the V/I converter, and (3) a reference voltage. It performs a monolithic signal conditioning of low-level inputs from a variety of transducers to control a standard 4 to 20 mA, two-wire current loop [7], and has a large range of applications [3].

### 10.1.3 Ratiometric Voltage Outputs

The stabilization of sensors supply voltage is usual but sometimes is not possible due to the high power consumption or additional costs. In these cases, it is desirable for the output of the sensor to be ratiometric with respect to its supply voltage. This output $V_{\text{OUT}}$ has to be permanently proportional to the supply voltage $V_S$, as generally expressed in the formula:

$$V_{\text{OUT}} = V_S \cdot TF \cdot CF \cdot V_{\text{IN}} \tag{10.9}$$

where $V_{\text{IN}}$ is the input signal. The transfer function TF and the correction function CF are assumed to be independent of $V_S$.

A series of sensors have been designed to fulfill this need (e.g., temperature sensors series AD2210x/Analog Devices). Autarkic ratiometric output stages are also available [5].

An important application of ratiometric sensors is the signal evaluation and supplying by means of ADCs (Section 9.8.2). The ADC voltage reference becomes the power supply source for the ratiometric sensor and its output is tied to the ADC input (Figure 10.5(a)). This step eliminates the need for a precision voltage reference in the sensor. However, it should be noted that the fact that a device is ratiometric does not preclude operating the part from a fixed reference voltage.

The low-cost, compact (SO8-package) ratiometric interface IC AM417 [8] belongs to a series of ratiometric devices [5]. It was developed for the conditioning of bridge signals in a very large temperature range: –40°C to +125°C and consists of the following three stages:

- The instrumentation amplifier IA, having a similar functionality as in the device AM410 (Figure 10.1), an input voltage range: 0 to 200 mV and a fixed gain of $G_{IA} = 10$;

- The ratiometric current source IB, particularly suitable for excitations of sensing elements in a constant current mode (0.5 to 1.25 mA). It is controlled by the op-amp OP, which is internally linked to the supply voltage pin VCC (5V, typically) via a 10:1 voltage divider (resistors 9R and R). With OP acting as a ratiometric control device, the output current is given by:

$$I_{IB} = \frac{V_S}{10 \cdot R_1} \tag{10.10}$$

and can be adjusted in the above-specified range by means of the resistor $R_1$.

- The output stage acts with the external transistor $T_1$ as a voltage amplifier that provides an output voltage range $V_{\mathrm{OUT}} = 0.5$ to 4.5V and can drive an output current of $I_{\mathrm{OUT}} \leq 11$ mA. The gain of the amplifier is:

$$G_{\mathrm{OUT}} = \frac{R_3 + R_4}{R_4} \tag{10.11}$$

Thus, the gain of the entire channel is:

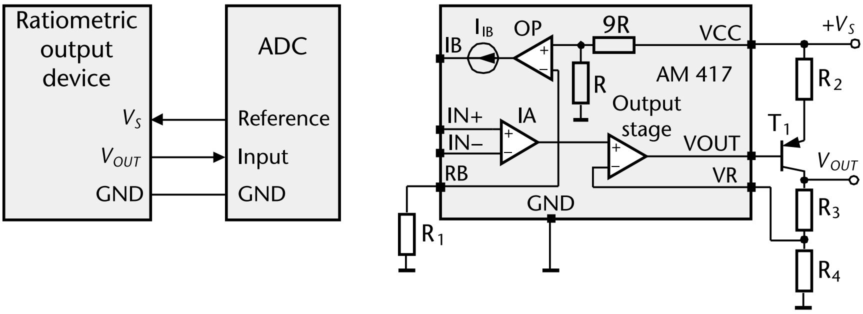
$$G_{SYS} = G_{IA} \cdot G_{\mathrm{OUT}} = 10 \cdot \frac{R_3 + R_4}{R_4} \tag{10.12}$$

A current limitation is integrated into the output stage. The maximum output current can be adjusted using the optional resistor $R_2$ according to the expression:

$$I_{\mathrm{OUTmax}} \approx \frac{0.38\ V}{R_2} \tag{10.13}$$

## 10.2 Output Drivers for Digital Inductive Sensors

According to the standards, inductive proximity sensors with switching output (also called inductive switches) must have a controlled semiconductor output switching



(a) (b)

**Figure 10.5** (a) Ratiometric output device in cooperation with an ADC, and (b) typical circuitry of the ratiometric output stage AM417 [5].

element (see the second attribute in the list with three essential characteristics, Section 2.1.1.1).

The spectrum of such switching outputs ranges from a discrete stage with a single transistor (bipolar, MOSFET, or DMOS), as exemplified in Figure 7.3 or 7.19, to a Darlington device containing a cascade of two or three transistors (Figure 7.15) and up to integrated output drivers (Figure 7.21).

Independent of the implemented version, standard requirements have to be fulfilled (Sections 1.3.7.1 and 2.1.1). The choice of a solution is only a question of two conflicting demands: hardware costs versus functional reliability, EMC immunity as well as additional features and intelligence, which are not specified in the standards.

### 10.2.1 Switched Inductive Loads, Voltage Clamps

The real loads of the inductive sensors are not pure resistive. The capacitive components of the real load connected to the switching sensor output—but predominantly its inductive components—are energy-storage elements. They should be considered as a memory that contains information stored in the past and permits history to play a role in determining the present and future states of the circuit.

The prime constituents of the clamp are the magnetic energy-storage element (load inductance $L_L$) supplying the additional energy (previously stored) and the diode providing short storage (charge) and long-decay time constants. These elements are considered below ideal elements.

Let us suppose an output with a resistive-inductive real load, with no stored initial energy in inductor, without clamping diode, and a square wave input signal $v_{\text{IN}}$ of the switching output transistor $\text{T}_{\text{OUT}}$ (Figure 10.6(a)).

At $t = 0$, when an input pulse is applied to the base, $\text{T}_{\text{OUT}}$ switches on during an ideal turn on time considered zero. The load is charged and the output $v_{\text{OUT}}$ now charges toward $V_{CC}$ with the time constant $\tau_1 = L_L/R_L$. The defining equation becomes

$$v_{\text{OUT}}(t) = V_{\text{CC}}\left(1 - e^{-t/\tau_1}\right) \tag{10.14}$$

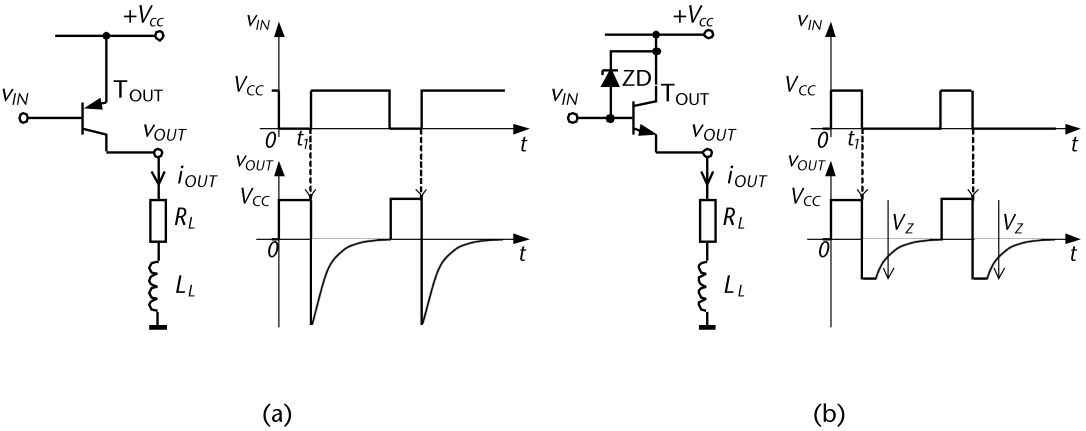


**Figure 10.6** (a) Switching of a resistive-inductive load (schematic and waveshape), and (b) the Zener diode (ZD) clamping mechanism (schematic and waveshape).

where the time constant has application-specific low values:

$$\tau_1 \leq \frac{1\text{ mH}}{100\Omega} = 10\ \mu\text{s} \tag{10.15}$$

Charging is virtually complete in $5\tau_1$. The transistor saturates and the current flowing through the load is constant:

$$I_{\text{OUT}} = \frac{V_{CC} - V_{CEsat}}{R_L} \approx \frac{V_{CC}}{R_L} \tag{10.16}$$

where $V_{CEsat}$ is the collector-emitter saturation voltage of $\text{T}_{\text{OUT}}$ ($\leq 0.1$ to 0.3V).

At $t = t_1$, when the pulse is removed and the base signal terminates, the transistor switches off. Supposing its conduction ceases immediately, the load current is almost instantaneously interrupted, the magnetic energy stored in the inductance induces a peak reverse voltage (back-emf) in the inductive load, and the output voltage $v_{\text{OUT}}$ makes a kick-back pulse. Correspondingly, this voltage becomes negative, and theoretically approaches the value of $-\infty$.

In the time interval $t > t_1$, the dissipated energy in the inductor's magnetic field recovers the output voltage toward the zero level along an exponential curve with a long completion time, determined by the inductance $L_L$ but also by the deep of $v_{\text{OUT}}$.

Breakdown-region operation during recovery generally has two hard effects. First, an excessive amount of time will have to be allotted for the recovery interval. Second, since the slow recovery is accompanied by very high collector voltage and high collector currents, the power dissipation will be large. If these exceed the transistor limitation, this can then be permanently damaged.

The negative glitch is highly dangerous for $\text{T}_{\text{OUT}}$. If, during flyback, the voltage across the transistor exceeds the reverse breakdown voltage, it can cause irreversible destruction of the transistor structure. In addition, it figures as a forbidden false impulse at the sensor output (see Section 2.1). To prevent this case, two solutions could be implemented:

1. To connect a free-wheeling diode (eventually in series with a resistor) in parallel to the last (reversed connection). During the time interval 0 to $t_1$, the diode is back-biased and remains OFF, and thus, the resistor is out of the circuit.

   After $t = t_1$, the diode conducts and clamps the negative output voltage on its forward-voltage. The resistor increases the peak flyback voltage and hence the deep of the negative clamping level, but advantageously reduces the recovery time. The diode stands high energetic stress and the recovery time until $v_{\text{OUT}}$ becomes zero is shorter. This historical method to fit relays with parallel diode is not popular by the sensor users.
2. A second possibility is to connect a Zener diode (ZD) in parallel to the BC junction of $\text{T}_{\text{OUT}}$ (Figure 10.6(b)). If $\text{T}_{\text{OUT}}$ switches off, a series circuit containing the back-biased Zener diode, BE junction of $\text{T}_{\text{OUT}}$ and the load results connected between $V_{CC}$ and GND. The negative glitch is cut on a negative value (equal to $V_{CC}$ minus Zener and BE voltages) and held on this level until

the induced voltage lowers the Zener voltage. Thereafter, $v_{OUT}$ drops along the exponential curve until it becomes zero. The clamped voltage prevents the transistor from destruction.

### 10.2.2 Output Drivers with Commercial Parts

A performant commercial output driver TDE1708DFT was already described in Section 7.4.2.

The current section follows this presentation raw with the device UC27132, called Smart Power Switch by the manufacturer Texas Instruments [9]. This device is designed to adapt inductive sensor schematics, which work at low supply voltage and deliver low power output signals to high supply voltage and high load current conditions.

#### 10.2.2.1 Smart Power Switch: Operation and Features

The circuit is the result of a from-customer-to-customer activity. It was specified for an optimal cost/performance ratio by a global sensor provider, designed and tested by the cooperation team: orderer–semiconductor manufacturer, and is available as a customer ASIC but also a catalog device. Essentially, there is a single chip with versatile usage. There are three packaged versions: low-side, high-side, and both sides, and every version is available for three temperature ranges: automotive range (–55°C … +125°C), industrial range (–40°C to +85°C), and commercial range (0°C to +70°C), respectively.

The monolithic ASIC for binary sensors outputs is realized using a high-voltage bipolar technology and consists of the following six stages (Figure 10.7):

- A robust 6V voltage reference, which can be supplied in a very large range: $V_{CC} = +8$ to +60V and provides the regulated voltage $V_{REF} = +6V \pm 0.2V$ based on an internal bandgap. Hence, the 6V reference has a low temperature dependence (less than 200 ppm/°C). The block is designed to bias the on-chip electronics and to power up external circuitry (preregulator with 8-mA current capability).
- The input comparator connected to the pin IN is a hysteresis comparator with low input current ($\leq 5\ \mu A$) and high gain, which fully switches with an analog or digital signal (0 to 6V max). Its ratiometric threshold is fixed (1/2 $V_{REF} = +3V$) and only a 5-mV overdrive is needed to switch the comparator.

  The hysteresis is adjustable in a large range by means of the resistor $R_{HYS}$ (0 to 100 kΩ) connected between the pin HYST and GND. The 100-kΩ resistor programs the smallest hysteresis of 30 mV (1%). For digital input signals, the hysteresis can be set to 900 mV (30% of 3V) connecting the pin HYST to GND.
- The logic stage is basically an AND gate, which combines the output of the comparator with an enable signal provided by the timing stage.
- The timing stage manages the time-related functions (Section 10.2.2.2) in tandem with the timing delay capacitor $C_{DEL}$ (1 to 47 nF, 10 nF typically.). If there is any fault situation, the block releases the logic unit and the output of the comparator directly controls the driver output stage. Under fault or

wanted-inhibit conditions, the AND gate is disabled and the output stage remains switched off.

The output stage, simplified and represented in Figure 10.7 by the transistors $T_1$ and $T_2$, is a composite PNP, NPN structure. It is a specially designed structure that keeps the drive current from the load current through the pin LS. The equivalent PNP Darlington transistor has equivalent open-collector (pin high-side (HS)) and equivalent open-emitter (pin low-side (LS)) terminals.

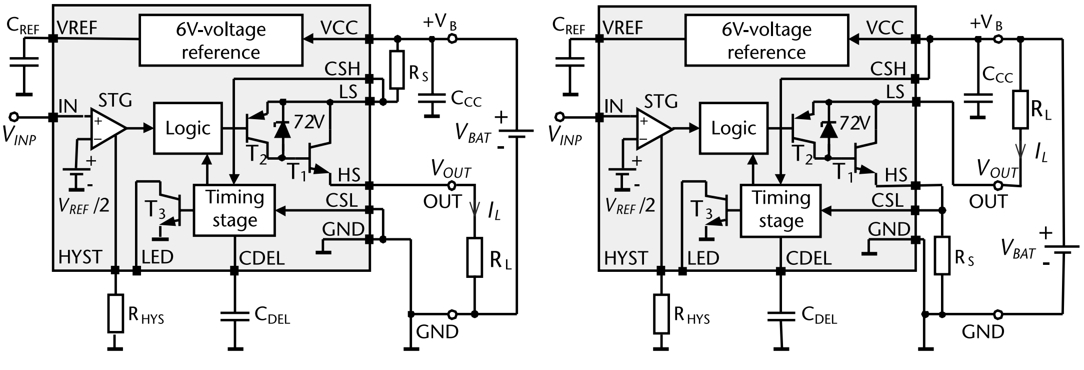
It should be mentioned that the chip has an integrated 72V power Zener diode as an internal clamping element (Section 10.2.1). This allows the output LS (in the LSD application, Figure 10.7(b)) to swing and clamp to $V_Z$ = 72V above GND when discharging an inductive load after an output current switching off. The clamp structure can discharge the 250- to 400-mA full load current.

Similarly, the Zener diode allows the output HS to safely swing and clamp $V_Z$ = 72V below VCC/LS when discharging an inductive load in a HSD application (see Figure 10.7(b)). The benefits of this smart solution are described in Section 10.2.1.

- The low-power switching transistor T3 is switched on by the timing stage synchronously with the disable signal going to the logic (fault situation). Its open collector is tied with the pin LED and provides two options: either to signalize the fault conditions to the front-end or to switch an LED ON/OFF, which is biased from VREF through a series resistor.

Designed and specified for a large supply voltage range ($V_{CC}$ = +8 to +60V), for high output currents (≤300 mA with internal clamping) and for extremely large temperature ranges with a low power dissipation (quiescent current 2.5 mA, typically), the device is a veritable output driver for sensor industrial applications. Additional technical data is as follows:

- Despite the used high voltage process ($V_{CC\max}$ = 80V) with larger structures, the chip area is reasonably maintained and the chip can also be placed in the cavity of an SOIC-8 package;



(a) (b)

**Figure 10.7** Smart Power Switch used as (a) high-side driver and (b) low-side driver.

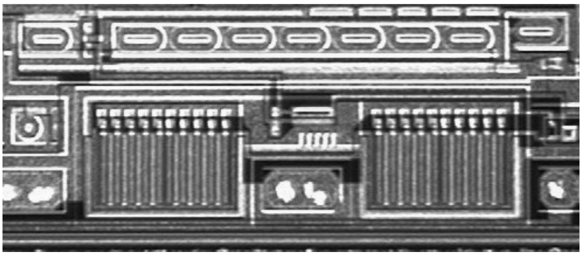
- Low voltage drop across conducting output transistor: $V_{HS\text{-}LS} \leq 1.2$ to 1.4V depending on the load current and temperature;
- Slew rates for OFF to ON or ON to OFF transitions: 25 and 15 V/$\mu$s, typically;
- Leakage current of the nonconducting output transistor: $\leq 5\ \mu$A;
- Capability of magnetizing and demagnetizing inductive loads up to 1 H without flyback voltage troubles;
- Capability of driving capacitive loads $\leq 1\ \mu$F, without short-circuit troubles;
- Current capability and leakage current of the LED output: $\leq 8$ mA/$\leq 5\ \mu$A.

The chip has an area that is a little larger than $3 \times 2$ mm$^2$ and has 14 regular pads (square shape: $100 \times 100\ \mu$m) (Section 7.3.1). The pads LS and HS have a double length. To reduce the current density in the bonding wires, a double bonding is performed at these places (Figure 10.8(b)).

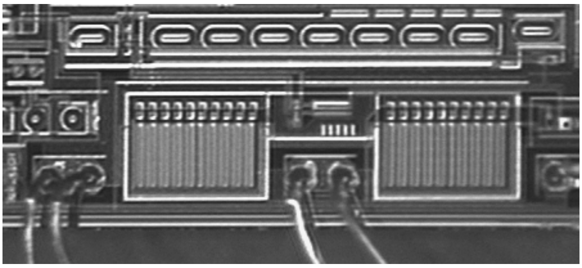
#### 10.2.2.2 Smart Power Switch: Protections and Timing Functions

Besides the remarkable main energetic features, the name "smart switch" is also earned by the versatile protection functions as well as fully integrated, one-capacitor adjusting timing:

a. The turn-on and turn-off delayed actions are globally defined in Section 1.3.6 and refer to the behavior of the switching output if the supply voltage is switched on and switched off, respectively. The operation of the Smart Power Switch totally complies with this standard definition.



(a)



(b)

**Figure 10.8** Photo cuts showing the chip area with the integrated output switching NPN transistor: (a) chip after probing (observe the probing needle marks on the connecting pads), (b) the same chip with bonding wires.

In order to prevent initial instability and oscillations of the output as long as the steady states are not achieved in the different sensor stages, a global gating time called: turn-on delay (TonD) is implemented. The output will be released only after this time (Figure 10.9). A short false output pulse (FOP) of the output voltage at the beginning of the TonD is possible and allowed by standards. Its duration is yet $(t_1 - t_0) \leq 1$ ms.

A similar instability could occur if the supply voltage is disconnected, reason to define and realize a second delay, namely the turn-off delay (ToffD). Usual time values are $TonD \leq 100$ ms and $ToffD \leq 10$ ms.

The Smart Power Switch has two additional attributes:

- Both delay times are linearly adjusted only with a single capacitor $C_{DEL}$, which is charged/discharged by two current source/sink stages between fixed-voltage thresholds 1.3 and 5.8V. As a result, the linear set equations are (typical values):

$$TonD = 1.1\ \frac{\text{ms}}{\text{nF}} \cdot C_{\text{DEL}}\ [\text{nF}] \text{ and } ToffD = 50\ \frac{\mu\text{s}}{\text{nF}} \cdot C_{\text{DEL}}\ [\text{nF}] \qquad (10.17)$$

- For some reason, it is possible that the supply voltage will have an accidental short interruption in industrial environments. This glitch could be considered a power switching off. Normally, the supply voltage immediately returns but a new TonD time is triggered, which is equivalent to a time extension of the interruption and is interpreted as a false output pulse. To prevent this situation, a waiting time after power switching off, called power interrupt period (PIP) is implemented. During this time interval, the device is supplied from power supply bypass capacitors and maintained in a standby mode. Hence, the TonD is retriggered only after an elapsed time of $t_{PIP} = 0.8$ ms, typically (Figure 10.9). If the glitch is shorter than $t_{PIP}$, the power interruption is ignored and output instantaneously restarts.

b. The overload and short-circuit protection (SCP) prevents the output transistors from irreversible damage in such cases. Section 2.1.1.5.5 defines these abnormal sensor working conditions and describes the immunity test to verify this protection, according to the standards.

The Smart Power Switch has a performant protection circuitry. This accomplishes a pulsing low-energy procedure that (1) does not produce thermal stress during the short circuit, (2) can protect the sensor for an infinitely long time, and (3) immediately returns to normal operation if the short circuit disappears or is removed. The procedure essentially corresponds to an astable multivibrator.

Depending on the chosen switching direction, a sense resistor $R_S$ (recommended value $R_S = 0.5\Omega$) is connected between the pins VCC and current sense high (CSH) or between GND and current sense low (CSL) (Figure 10.7). The load current flows through this part and causes a voltage that is measured by the inputs CSH or CSL. If the sense voltage exceeds 150 mV (typical), the pulsing protection function is activated and turns off the switching transistors.

After an adjusted recovery time $t_P$, in which $C_{DEL}$ is charged, the transistors are turned on for a very short $t_{SC}$ time in order to check if the short circuit persists. This time $t_{SC}$, in which $C_{DEL}$ is discharged, is short enough to prevent the destruction of the switching transistors. This is also a useful inhibit function since brief glitches on the outputs or the charging of sensor line capacitances should not trigger the short-circuit protection. The output transistors operate now as current source/sink with a current capability between 300 and 400 mA.

If the failure persists, the transistors are turned off again. The multivibrator is continuously running until the mishandling ends or it was removed.

At the same time, the capacitor $C_{DEL}$ also determines the timing of the SCP according to the following formulas and typical values:

$$t_{SC} = 7.5\ \frac{\mu\text{s}}{\text{nF}} \cdot \text{C}_{\text{DEL}}\ [\text{nF}] \text{ and } t_P = 1\ \frac{\text{ms}}{\text{nF}} \cdot \text{C}_{\text{DEL}}\ [\text{nF}] \tag{10.18}$$

The ratio between the recovery time $t_P$ and checking time $t_{SC}$ is great enough:

$$t_P : t_{SC} \approx 133 : 1 \tag{10.19}$$

to assure the ASIC protection (low-duty cycle mode).

### 10.2.3 Monolithic Integrated Output Drivers in ASICs

The most valuable implementation of the output driver is to integrate this high-voltage, high-current stage as a back-end on the same chip with the low-voltage, low-current sensor front-end electronics. The monolithic solution allows back-end to get more information and to react with increased intelligence. On the other hand, to balance the different energetic conditions on the same chip, to prevent a parasitic back-end influence on the sensitive front-end are high design skills.

Nonetheless, such ASICs are today reality and users report full satisfaction. The sensor manufacturer Contrinex AG / Switzerland (www.contrinex.com) presented a new inductive sensors family called MiniDist® [10]. This fulfills the two conflicting demands: very small sizes (external housing diameter of 4 mm or M5 threaded tube

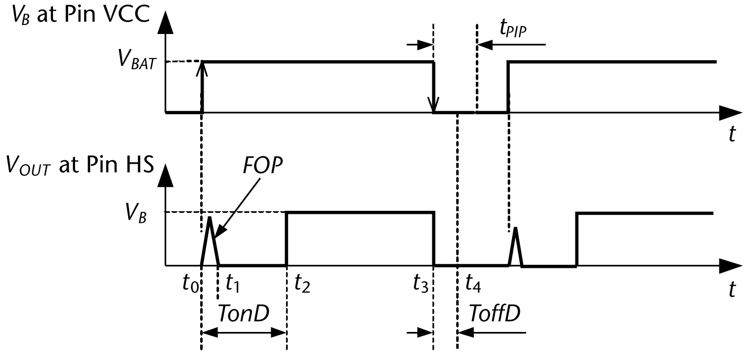


**Figure 10.9** Definitions of the TonD and ToffD.

with a length of 25 mm) in parallel with very high operating distance equal to 2.5 mm, which is three times $s_n$ (Section 2.1.1.2) and increased reliability. Contrinex affirms that this so-called "quantum leap" in sensor technology was possible only through the development and use of a system on chip ASIC. The Condist® ASIC is a single chip (Figure 10.10) that successfully combines front-end demands with back-end features. Realized in a 0.7-$\mu$m CMOS process, the chip can be supplied with a 10 to 30 VDC as standard and can switch load currents up to 200 mA. Protection functions are also inside.

In the monolithic implementation, the switching transistors of the output stages are realized by connecting a large number of regular transistors contained in the process specification in parallel, strictly speaking the same transistors as for the front-end. Computation of the necessary number, and thus of the required area on the chip, depends on the features of the elementary transistor, on the accepted current densities in the structures, and in the copper conductive paths, as well as on their resistance.

Finally, such decisions determine the power dissipation and thus the chip heating as well as the electrical data, as voltage drop at the switched on output, jump current capability for capacitive loads, and so forth. The placement of the semiconductor regions, pads, and conductive lines plays a massive role in the layout design (Figure 10.11).

An orientating framework about the chip area consumption to integrate a bidirectional output stage with different HSD and LSD transistors for 250 mA, a proof

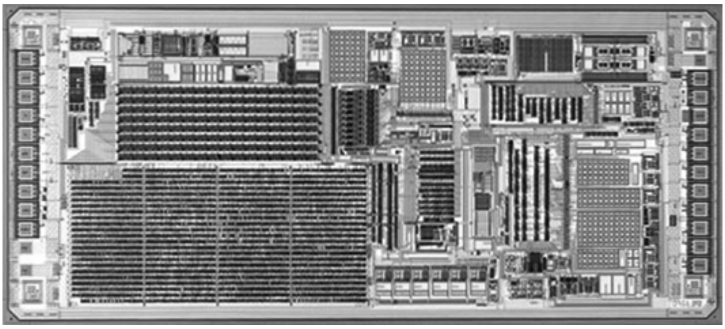
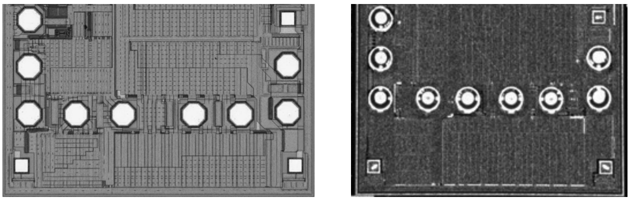


**Figure 10.10** The Condist® ASIC layout [10].



(a) (b)

**Figure 10.11** Chip area with integrated NMOS and PMOS output transistors (36V/250 mA) (a) Layout picture and (b) chip photo.

voltage of 35V, and using a standard 0.6–$\mu$m CMOS process that fulfils industrial demands should be

- For the PMOS architecture (HSD output): a total channel width of $W \geq 30.000$ $\mu$m (channel length $L = 6$ $\mu$m);
- For the NMOS architecture (LSD output): a total channel width of $W \geq 10.000$ $\mu$m (channel length $L = 6$ $\mu$m).

Adding the area needed for connection strips, pads, chip border, and so forth, the estimated total chip area results between 1.5 and 2 mm$^2$.

In monolithic implementations, a small fraction of the load current flowing through the main transistor can be deviated and measured to detect an overload. Another overload detection approach is to indirectly measure the temperature in the hot region of the integrated structure and to trigger a pulsing function as mentioned above [11].

## References

[1] https://www.analog-micro.com/_pages/ics/am401/am401_data_sheet.pdf, data sheet, AM401, Industrial Voltage Amplifier IC, Analog Microelectronics GmbH/Germany, 2006.

[2] Gray, P. R., P. J. Hurst, S. H. Lewis, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, Fifth Edition, New York: John Wiley & Sons, 2009.

[3] Nawrocki, W., *Measurement Systems and Sensors*, Second Edition, Norwood, MA: Artech House, 2016.

[4] https://www.analog-micro.com/_pages/ics/am462/am462_data_sheet.pdf, data sheet, AM462, Industrial V/I Converter and Protector IC, Analog Microelectronics GmbH/Germany, 2007.

[5] https://www.analog-micro.com/en/products/ics/, Integrated Circuits, Analog Microelectronics GmbH/Germany, 2018.

[6] www.analog.com/media/en/technical.../data-sheets/AD694.pdf, data sheet, AD694, 4–20 mA Transmitter, Analog Devices, 2002.

[7] www.analog.com/media/en/technical.../data-sheets/AD693.pdf, data sheet, AD693, Loop-Powered 4–20 mA Sensor Transmitter, Analog Devices, 2002.

[8] https://www.analog-micro.com/_pages/ics/am417/am417_data_sheet.pdf, data sheet, AM417–Ratiometric Instrumentation Amplifier with Adjustable Output Stage, Analog Microelectronics GmbH/Germany, 2008.

[9] http://www.ti.com/lit/ds/symlink/uc37133.pdf, data sheet, UC37133, Smart Power Switch, Texas Instruments, 1999.

[10] Heimlicher, P., *MiniDist®-Miniaturschalter D4 und M5 mit großem Schaltabstand*, Contrinex AG, Presentation, Fachpressetage, Bruchsal, Germany, 2005.

[11] Konishi, Y. (Omron Corporation), Sensor Output IC and Sensor Device, U.S. Patent 8,629,710 and European Patent EP 2367288B1, October, 2010.

CHAPTER 11

# Inductive Sensors: Power Supply and Sensor Protections

## 11.1 Power Supply Circuits

The regulated power supply is quite an unavoidable function in every inductive sensor. The ideal regulated power supply (voltage regulator) is an electronic circuit (see the stage PSP in Figure 7.1) designed to provide a predetermined DC voltage $V_{CC}$ that is independent of the current $I_{CC}$ drawn from $V_{CC}$, of the temperature, and also of any variations of the supply voltage $V_B$.

The $V_B$ can be regulated (e.g., a + 24V DC local sensor supply network) or unregulated provided by machine tools supply units, which contain a voltage transformer, a rectifier, and a filter. Regarding the value of $V_B$, the standards state large operational voltage ranges (Table 2.3).

There are three reasons why a regulated power supply should be used in an inductive sensor:

1. *The good load regulation.* The output voltage $V_{CC}$ should remain constant as the load varies. The load current represents the sensor electronics current consumption and is permanently variable.
2. *The good line regulation.* The output voltage $V_{CC}$ should be as far as possible constant as the input voltage $V_B$ varies.
3. *The temperature drift of the* $V_{CC}$. In particular, because semiconductor devices are used, this should be low enough not to influence the sensor electronics.

The inductive sensors for industrial applications have to fulfill the specific standard demands. Accordingly, only a single *unipolar supply voltage* ($+V_{CC}$ to GND) is available for the regulated power supply. Dual supply topologies ($+V_{DD}$, GND and $-V_{SS}$) are generally not accepted in most industrial applications.

The sensor supply voltage lies in the range $V_{CC}$ = 2V to 30V and the power consumption is relatively low (below 500 mW). This means the load current $I_{CC}$ is usually below 100 mA.

The voltage regulators groupings [1] consist of two main categories: (1) those with discrete components, and (2) those with integrated circuits. Both can be subdivided further into (1) linear regulators and (2) shunt regulators. Today, sensors are almost exclusively constructed using integrated regulators. Another classification distinguishes low-accuracy, medium-accuracy, and high-accuracy regulators, also called references.

## 11.1.1 Series Voltage Regulators/References

### 11.1.1.1 Basics of Series Voltage Regulators

The major component of a series voltage regulator is the pass element. This is connected in series with the supply voltage source $V_B$ and with the load, namely the sensor electronics to be supplied (Figure 11.1(a)). This series topology of the main current path gives the name of the regulator. The pass element must absorb a part $\Delta V_{PE}$ of the voltage $V_B$ (static regulation) as well as temporal $V_B$ fluctuations (dynamic regulation) to ensure a constant, ripple-free output voltage $V_{CC}$:

$$V_{CC} = V_B - \Delta V_{PE} = \text{constant} \tag{11.1}$$

To do that, the pass element is included in a regulating loop (see Figure 8.9). A differential amplifier compares the information about the output $V_{CC}$ with a reference voltage and controls the pass element to fulfill (11.1). When these magnitudes are equal, the control is held firm, as long as a change of $V_B$ or $I_{CC}$ does not appear.

The classical implementation in Figure 11.1(b) is a voltage-series feedback circuit that realizes the three desiderata mentioned above, and also reduces the ripple voltage. Assuming that the voltage gain of the emitter follower $T_1$ (pass element) is approximately unity, then $V_{CC} \approx V_O$ and:

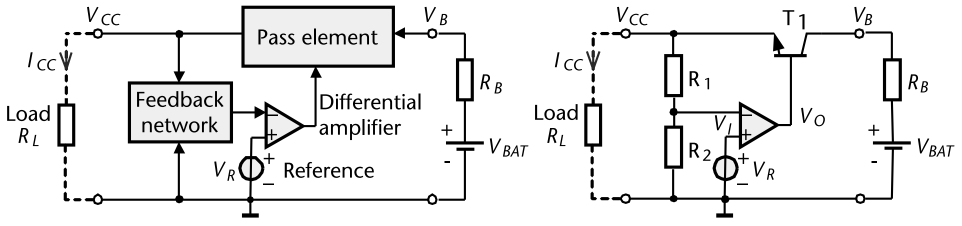
$$V_O = A_V V_I = A_V\left(V_R - \beta V_{CC}\right) \approx V_{CC} \tag{11.2}$$

where the feedback factor is:

$$\beta \equiv \frac{R_2}{R_1 + R_2} \tag{11.3}$$

From (11.2) it follows for $A_V >> 1$:

$$V_O = V_R \frac{A_V}{1 + \beta A_V} \approx \frac{V_R}{\beta} = V_R \frac{R_1 + R_2}{R_2} \tag{11.4}$$



(a) (b)

**Figure 11.1** Series voltage regulator: (a) functional block schematic, and (b) circuit diagram of the classical implementation.

The output voltage can be changed by varying the feedback factor $\beta$. The emitter follower with $T_1$ is used to provide current gain, because the current delivered by the differential amplifier is usually not sufficient.

Since the output voltage $V_{CC}$ depends on the input voltage $V_B$, load current $I_{CC}$, and temperature $T$, then the change $\Delta V_{CC}$ in output voltage of the power supply can be expressed as follows:

$$\Delta V_{CC} = \frac{\partial V_{CC}}{\partial V_B}\Delta V_B + \frac{\partial V_{CC}}{\partial I_{CC}}\Delta I_{CC} + \frac{\partial V_{CC}}{\partial T}\Delta T \tag{11.5}$$

or

$$\Delta V_{CC} = S_V \Delta V_B + R_o \Delta I_{CC} + S_T \Delta T \tag{11.6}$$

The three coefficients are defined as:

1. Input regulation factor, specifying the line regulation:

$$S_V = \frac{\Delta V_{CC}}{\Delta V_B} \text{ at } \Delta I_{CC} = 0,\ \Delta T = 0 \tag{11.7}$$

2. Output resistance, specifying the load regulation:

$$R_0 = \frac{\Delta V_{CC}}{\Delta I_{CC}} \text{ at } \Delta V_B = 0,\ \Delta T = 0 \tag{11.8}$$

3. Temperature coefficient, characterizing the temperature dependence:

$$S_T = \frac{\Delta V_{CC}}{\Delta T} \text{ at } \Delta V_B = 0,\ \Delta I_{CC} = 0 \tag{11.9}$$

The smaller the value of the three coefficients, the better the regulation of the power supply. In addition to these coefficients, two practice-oriented features are used to specify the voltage regulators:

1. Power supply rejection ratio PSRR, expressed in decibels:

$$PSRR = 20\log\frac{\Delta V_B}{\Delta V_{CC}}\ [\text{dB}] \text{ at } \Delta I_{CC} = 0,\ \Delta T = 0 \tag{11.10}$$

2. Related temperature coefficient TC, expressed in parts per million per °C:

$$TC = \frac{10^6}{V_{CC@25°C}} \cdot \frac{(\mathrm{V}_{CC})_{T2} - (\mathrm{V}_{CC})_{T1}}{T_2 - T_1}\ [\text{ppm/°C}] \text{ at } \Delta V_B = 0,\ \Delta I_{CC} = 0 \tag{11.11}$$

where $V_{CC@25°C}$, $(V_{CC})_{T1}$, $(V_{CC})_{T2}$ are the $V_{CC}$ voltages at room temperature and at temperatures $T_1$ and $T_2$.

The greater the value of $PSRR$ and the smaller the value of $TC$, the better the regulation of the power supply.

Note that the schematic in Figure 11.1(b) was used in the past to construct discrete component regulators, using an op-amp (e.g., the classical LM741 [2]) and a Zener diode classically biased from $V_B$ or $V_{CC}$ through a resistor to generate the reference $V_R$.

An integrated circuit LM723, which belongs beside $\mu$A741 (see above), NE555 (see Section 8.4.3), and others in the pioneering times of integrated circuits but is still available [3], has an architecture similar to those in Figure 11.1(b). With improved stages (temperature compensated and buffered Zener diode, current limiting circuit, etc.), LM723 can be used with its internal pass transistor or can drive an external transistor, to increase the load current.

With the explosion of microelectronics it has become technically and economically feasible to incorporate all components and also connection lines on a monolithic chip. All benefits of ICs are thus obtained: excellent performance, small package size, easy to use, low cost, and high reliability. The result is the monolithic regulators with their inestimable market volume.

#### 11.1.1.2 Monolithic Regulators

Typically, monolithic regulators are compact, standard three-terminal voltage systems (Figure 11.2(a)).

Using monolithic regulators, it is possible to distribute unregulated or low-accuracy regulated voltages through electronic equipment and provide pure regulation locally. Among the advantages of this approach are greater flexibility in voltage levels, regulation for individual stages, and improved isolation and decoupling of these stages.

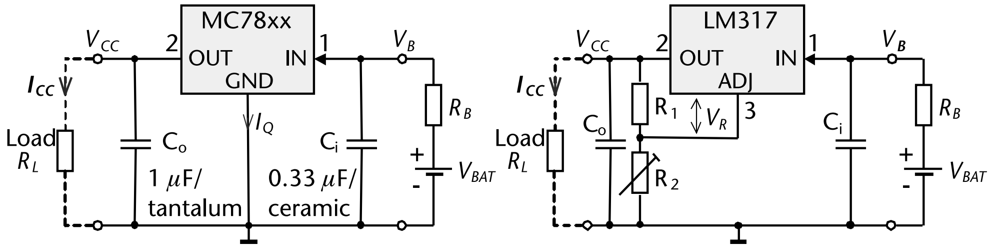
Monolithic regulators are available in a multitude of performance levels:

- Fixed (usually 5, 6, 8, 12, 15, 18, or 24V) or variable output voltage;
- Positive or negative output voltage;
- High output current (>1A);
- High output voltage (>24V);
- Single or dual (±) outputs, etc;
- Guide coefficient values: $S_V = 0.003$, $R_o = 30\ \text{m}\Omega$, $S_T = -1\ \text{mV/°C}$.

A conventional example of monolithic regulator represents the LM78xx series (three-terminal, positive fixed-voltage regulator) [4].

The standard application (Figure 11.2(a)) shows a minimized user complexity, which is finally reduced to two decoupling capacitors. The input capacitor $C_I$ is required to cancel inductive effects associated with long power-distribution wires. The output capacitor $C_O$ improves the transient response.

These devices, requiring no adjustment, have an output preset by the manufacturer to an industry standard voltage (see the list above). A small disadvantage results in the need of a minimum of 2V between input and output. This could lead



(a) (b)

**Figure 11.2** (a) Three-terminal, fixed-voltage monolithic regulator LM78xx, and (b) three-terminal, adjustable, monolithic regulator LM317.

to a limitation of the load current even when such regulators are capable of output currents in excess of 1.0A. The devices have internal short-circuit protection with current limitation and thermal shutdown and output-transistor safe operating area protection. The regulators are available in a large variety of packages from the huge body TO-3 down to SOT-223 or SO-3 [4].

The level of complexity can be appreciated by examining the circuit diagram in [4]. Essentially, the circuit consists of a Zener diode buffered by an emitter-follower, a differential amplifier having the design similarity to the LM741 [2], a resistor divider as in Figure 11.1(b), and specific protection schematics.

The pass element ($T_1$ in Figure 11.1(b)) is a Darlington pair of two NPN transistors, and that is the explanation why the voltage drop has to be ≥2V.

Using three-terminal adjustable regulators and two additional resistors, the regulated output voltage can be varied in large ranges. As an example, the LM317 device in Figure 11.2(b) is an adjustable positive-voltage regulator capable of providing more than 1.5A over an output-voltage range of 1.25 to 37V [5]. It includes current limiting, thermal overload protection, and safe operating area protection. Overload protection remains functional even if the adjust terminal ADJ is disconnected. The needed output voltage can be calculated using the equation:

$$V_{CC} \approx V_R \frac{R_1 + R_2}{R_1} = \frac{R_1 + R_2}{R_1} \cdot 1.25\text{V} \tag{11.12}$$

Finally, it is important to note the existence of a complementary family of three-terminal, negative fixed-voltage monolithic regulators LM79xx with similar features [6].

### 11.1.1.3 Low-Dropout Regulators

The low-dropout (LDO) voltage regulators are a second type of linear regulators. The difference between standard regulators and LDOs is in the pass element and the amount of headroom, or minimum dropout voltage, required to maintain a regulated output voltage.

For standard regulators, the pass element is a Darlington output stage (see the description of LM78xx in Section 11.1.1.2). Due to the two BE junctions in the series, standard linear regulators have voltage drops as high as 2V, which are acceptable for applications with large input-to-output voltage differences.

In contrast, applications with low input-to-output voltage differences require the use of an LDO to achieve the lower dropout voltage. The LDOs use an N-channel or P-channel FET pass element and can have dropout voltages less than 100 mV [7]. Except for the the pass element, the circuit diagram in Figure 11.3(a) is identical to those in Figure 11.1(b). The error amplifier drives the pass elements' gate to the appropriate operation point to satisfy the operation law (11.1). Under steady-state operating conditions, the LDO behaves like a resistor. However, operating conditions are never static; therefore, feedback is necessary to adapt the LDO's effective resistance.

The TPS769xx family of LDOs [8] features low-dropout voltages and ultralow quiescent current compared to conventional regulators. LDOs are offered in a five-terminal small outline SOT-23 package.

The TPS769xx is available in fixed-voltage versions: 1.2, 1.5, 1.8, 2.5, 2.7, 2.8, 3.0, 3.3, and 5V, and as a variable version (programmable output range from 1.2 to 5.5V). The most used input voltage is 5V. A 3.3V application of the LDO is shown in Figure 7.24 (see the LVDR-stage).

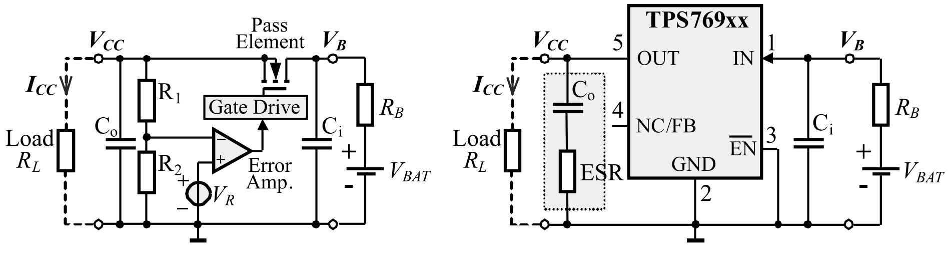
Because the PMOS pass element behaves as a low-value resistor, the dropout voltage is very low, typically 71 mV at 100 mA of load current, and is directly proportional to the load current.

Since the PMOS pass element is a voltage-driven device, the quiescent current is ultralow (28 $\mu$A maximum) and is stable over the entire range of output load current (0 to 100 mA). The TPS769xx also features a logic-enabled sleep mode to shut down the regulator, reducing quiescent current to 1 $\mu$A typical at room temperature.

The data sheet recommends an attentive choice of the output capacitor $C_O$ (Figure 11.3(b)). Beside its high capacitance, the value of its equivalent series resistance (ESR) should be as low as possible to ensure stability. A capacitor selection list is available in [8].

## 11.1.2 Shunt Voltage Regulators

Shunt voltage regulators are not widely used in many applications to provide the main regulation because they offer a low efficiency level. A maximum current is



(a) (b)

**Figure 11.3** LDO regulators: (a) block diagram in its most basic form with a MOS FET as pass element, and (b) LDO regulator application with TPS769xx.

drawn from the source regardless of the load current (i.e., even when there is no load current).

Despite their significantly narrower spread in electronics, the shunt regulators have a high relevance for inductive sensors, more exactly for the two-wire versions (Table 1.3) and that is why we present them in this section.

The basic operation schematic of a shunt regulator is a series consisting of a voltage source, a resistor, and the shunt regulator in parallel with the load.

In order to keep the voltage across the load constant, a level of current must be drawn through the series resistor to maintain the required voltage across the load. The load will take some current and the remaining current is drawn by the shunt voltage regulator. The circuit is designed so that at maximum load current, the shunt regulator draws virtually no current and at minimum load current, the shunt regulator passes the full current.

A common form of shunt regulator is the simple Zener diode. Once its small minimum current is exceeded, the Zener diode maintains an almost constant voltage across its terminals. The series resistor drops the voltage from the source to the Zener diode and load. Any variations in load current do not affect the voltage across the Zener diode. It takes up the current variations required to ensure the correct drop across the series resistor.

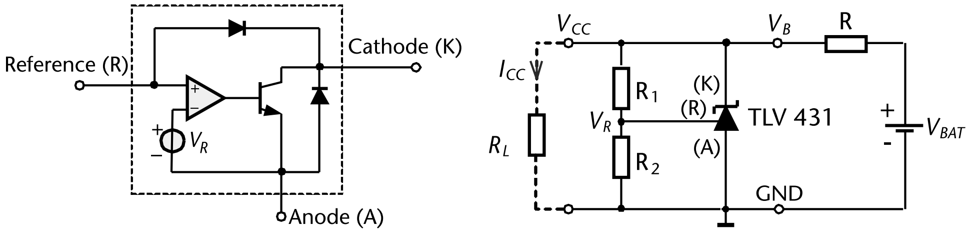
The Zener diode must be capable of dissipating the power from the maximum amount of current it is likely to handle. Thus the total maximum current that will be passed by the diode is the load current plus an allowance for current to maintain the reference voltage when the load is taking its maximum current.

It should also be noted that for the shunt regulator circuit, the series resistance is comprised of the series resistor value plus any source resistance. In most cases, the value of the series resistor will dominate and the source resistance can be ignored.

This elementary shunt regulator has many applications in the inner workings of electronic circuits as medium-accuracy voltage reference and was presented in previous chapters (e.g., Figure 7.3).

The basic shunt voltage regulator above does not have any feedback (i.e., it runs in an open-loop manner and its performance is sufficient for many applications).

Much higher levels of performance can be achieved by providing feedback based on the output voltage and feeding this back into the system to ensure that the required output voltage is accurately maintained. Using such feedback, the output voltage is sensed and the voltage is compared to a reference (Figure 11.4(a)).



(a) (b)

**Figure 11.4** Shunt voltage regulator: (a) representative block diagram, and (b) typical application circuit with TLV431.

The level of the shunt current is then altered to return the output voltage to the required level.

The worldwide established monolithic shunt voltage regulator is based on the 431-system architecture and is manufactured by several semiconductor companies (Texas Instruments, SEMTECH, ON Semiconductor, etc.) having different ID prefixes.

For example, the TI device TLV431 [9] is available in several performance levels:

- Low-voltage operation with a reference of $V_R$ = 1.24V and correspondingly adjustable output voltage $V_{CC}$ = 1.24 to 6V;
- Reference voltage tolerances: 0.5% or 1% or 1.5%;
- Typical temperature drift: 4 mV at (0°C to 70°C), 6 mV at (−40°C to 85°C), and 11 mV at (−40°C to 125°C);
- Low operational cathode current: 80 $\mu$A, typically;
- Typical output impedance: 0.25Ω;
- A multitude of packages: SOT-23-5, SOT-89, etc.

A high-voltage, high current version, SC431, with comparable features but with $V_{CC}$ = 2.5 to 3V and a current up to 150 mA is offered by SEMTECH / USA (www.semtech.com);

The typical application in Figure 11.4(b) shows the possibility to adjust the output voltage using a voltage divider (resistors $R_1$ and $R_2$), according to the formula:

$$V_{CC} \approx V_R \frac{R_1 + R_2}{R_1} = \frac{R_1 + R_2}{R_1} \cdot 1.24\text{V} \tag{11.13}$$

In addition, the device has two test facilities, which make it feasible for sensor two-wire versions:

1. Reference terminal $V_R$ bridged with the cathode and no other resistors provides $V_{CC} = V_R$ = 1.24V;
2. Reference terminal $V_R$ bridged with the anode and no other resistors provides OFF state of the device.

Regarding the functionality of a two-wire switching sensor (Section 1.3.7.1, Table 1.3), it is equivalent to an electrical light switch, is quite popular, and it finds universal use, especially in Japan and the Asian market due to its simplicity and lower installation costs.

The application consists of a series circuit: inductive sensor IPS, application load $R_L$ (relay, digital port input, etc.), and power supply source $V_{BAT}$ (Table 1.3). This signal chain has two states:

1. If the sensor (NO version) is actuated, the internal switching driver SW is active and the load current flows from the battery through the load and through the sensor. The voltage drop between $+V_B$ and $-V_B$ should be as low as possible. Zero volts is impossible because the sensor electronics needs a

supply voltage. Thus, the sensor electronics has to regulate the voltage drop to a minimum value (e.g., 2.5V for $I_L \leq 100$ mA).

2. Contrarily, if the sensor (NO version) is nonactuated, the internal switching driver SW is open and the current flow should be interrupted. In reality, it is necessary to maintain a residual current in order to supply the sensor with power. Competitive two-wire sensors require such cutoff currents below 1 mA. The voltage drop on the load is very small and the sensor sees the entire battery voltage.

This functionality can be optimally realized replacing the internal switching driver SW by an internal shunt regulator device (SRD) (Figure 11.5) accomplished by the voltage divider (resistors $R_1$ and $R_2$), and a low-power toggle-switch TSW. This is controlled by the sensor electronics via low-power output driver ODV as follows:

- If the NO sensor is actuated, the resistor $R_1$ is connected to the reference terminal of the SRD. Hence, SRD is normally biased by $R_1$ and $R_2$ and regulates the voltage drop between $+V_B$ and $-V_B$ on the wanted value. The sensor state condition 1 (see above) is totally fulfilled.
- If the NO sensor is released, $R_1$ is open and $R_2$ is short-circuited. As mentioned above, the SRD switches in its OFF state and interrupts the load current. The states corresponds to condition 2 above. The residual current of the sensor electronics represents the cutoff current.

If the supply voltage $V_{BAT}$ is applied, the circuit needs to start up with a specific default state, in which an initial sensor supply is performed. The sensor actuation conditions are checked and the switch TSW is set, correspondingly.

## 11.2 Standard and Supplimentary Sensor Protection Functions

Practice-oriented demands of the industrial applications and compliance with specific standards require sensors that are protected against electromagnetic (EMC) disturbances and also electrically protected against accidental wrong use.

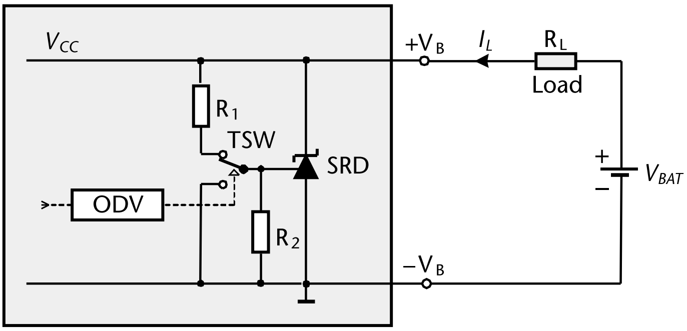


**Figure 11.5** Implementation of the SRD as an output driver of a two-wire inductive proximity sensor (see Table 1.3).

The first category of parasitic influences and the corresponding EMC requirements are treated in detail in Section 2.2.

The present section deals with electrical protections. Some of these concerns the sensor output (overcurrent, overvoltage, etc.) and are analyzed in Chapter 10.

The second class of electrical protections refers to the inductive sensor as a unitary system and are discussed in the next sections.

### 11.2.1 Open Wire Protection

Open wire protection (OWP) performs the task of preventing faulty sensor functions when an accidental interruption of one or more sensor wires occurs. In these cases, parasitical supply paths, and hence, faulty switching of the load, must be avoided.

An interruption of the positive or of the output line of an inductive sensor having a high-side driver (HSD) configuration (Figure 11.6(a)) carries to an intrinsic protection: the sensor or the load becomes currentless. The same thing happens if the ground or the output line of an inductive sensor having a low-side driver (LSD) configuration (Figure 11.6(b)) is interrupted.

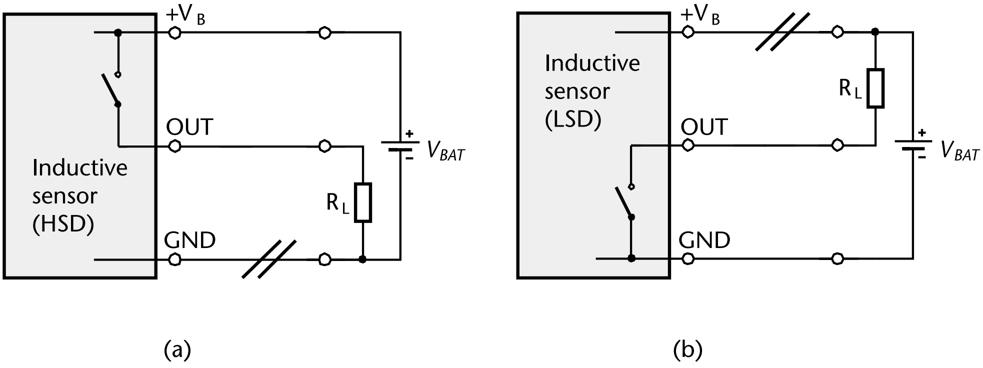
In contrast, if the ground line of an HSD or the plus line of an LSD is broken, parasitic currents paths through the two remaining intact wires (i.e., plus and output or output and ground) could appear. The load $R_L$ could be erroneously actuated.

The broken wire protection should avoid this current flow out of sensor output or in the sensor output. Protection implementations are very different; however, they could be classified into two categories:

1. Passive measures, consisting in the insertion of a current limiting element (e.g., resistor) or a unidirectional part (e.g., diode). Unfortunately, they downgrade the sensor features (higher voltage drop and power dissipation, etc.).
2. Active circuits, which recognize a disconnected sensor terminal and instantaneously disable the output in this case.

### 11.2.2 Reversed Polarity Protection

Corresponding to up-to-date market demands, inductive sensors should feature internal reverse polarity protection RPP. The RPP circuitry protects the device



(a) (b)

**Figure 11.6** Open-wire protected fault conditions: (a) broken GND wire for high-side driver configuration, and (b) broken $+V_B$ wire for low-side driver configuration.

against accidental reverse polarity connections to the terminals +$V_B$, OUT and GND. Protection means prevention of sensor malfunction or sensor damage as a result of reversed connections to the supply source and/or load.

For a system having three terminals, the number of connection possibilities is mathematically expressed by the factorial function: 3! = 6.

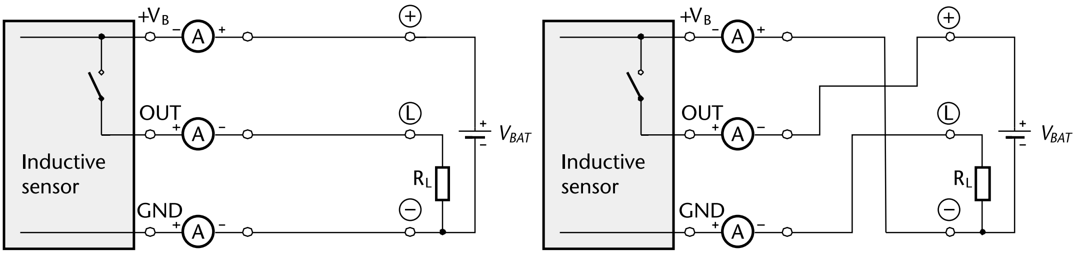
Figures 11.7 and 11.8 show all the possible connection combinations for high-side driver configuration and low-side driver configurations, respectively.

There is one correct configuration and five reverse polarity protected fault conditions, which can be subclassified in simple or multiple reverse polarities. The need is to generally have reverse currents equal to zero or not higher than a few milliamperes, currents that are not dangerous for sensor electronics.

The protection against plus and minus polarity reversal (reverse case number 1 in Figures 11.7 and 11.8) is a minimal requirement of a sensor for industrial applications.

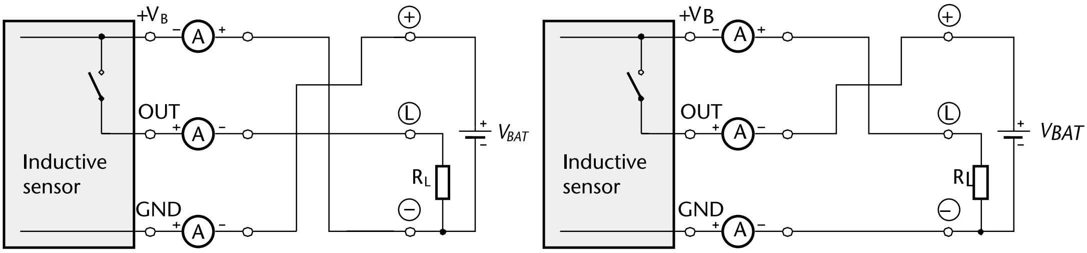
0. Right Polarity

3. Multiple Reverse Polarity



1. Simple Reverse Polarity

4. Simple Reverse Polarity



2. Simple Reverse Polarity

5. Multiple Reverse Polarity

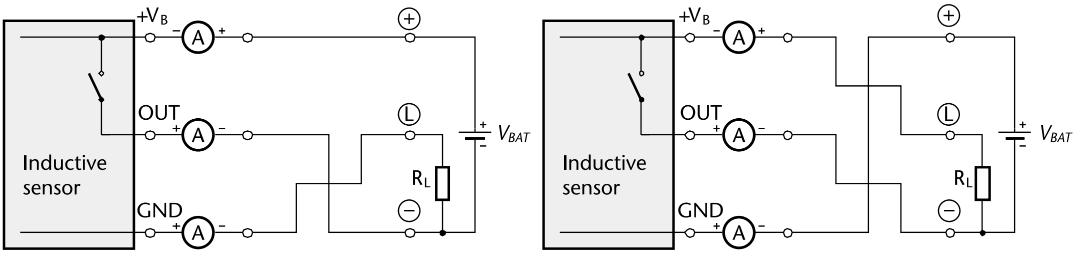
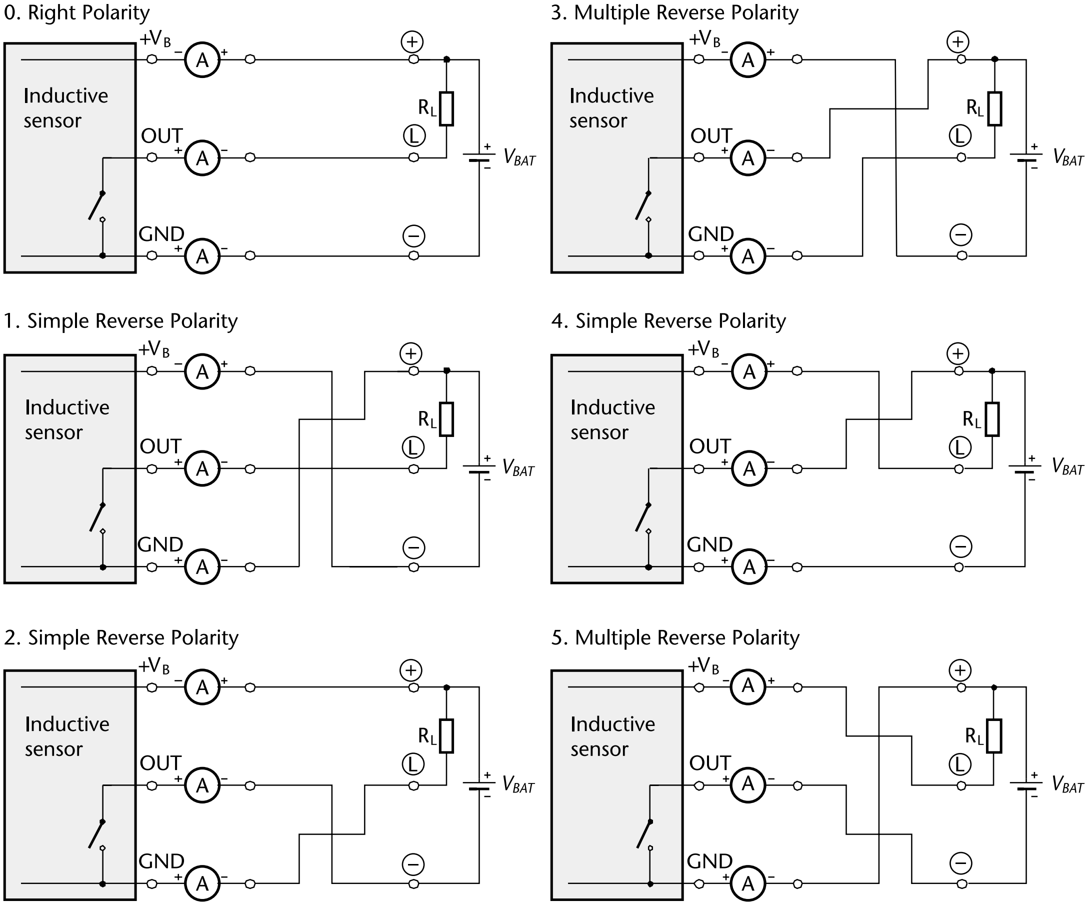


**Figure 11.7** One correct polarity use and five reverse polarity protected fault conditions for high-side driver sensor configuration [10].



0. Right Polarity

1. Simple Reverse Polarity

2. Simple Reverse Polarity

3. Multiple Reverse Polarity

4. Simple Reverse Polarity

5. Multiple Reverse Polarity

**Figure 11.8** One correct polarity use and five reverse polarity protected fault conditions for low-side driver sensor configuration [10].

Traditional protection methods use more diodes with the same consequences as those of the OWP protections. Full monolithic integration of sensor electronics offers versatility and a place for intelligent solutions.

To give a practical validated example from the author's field of experience, it is possible to integrate a low-current rectifying bridge, which is operating only during the sensor start-up (time interval TonD in Figure 10.9), provides a right sensor supply allowing a polarity check. For positive results, this technical aid is disconnected and replaced by high-current paths without additional dissipative protection parts. In contrast, if a reverse polarity is identified, the terminals are reversibly locked.

### 11.2.3 Protection against High-Energetic Pulses (Surge)

The causes of these high-energetic induced pulses in the sensor and the standard method and equipment to test and certify the resilience against such impulse voltages is treated in Section 2.2.4.

The highly increased voltage levels may be induced between sensor wires (supply and/or outputs) when a lightning strike occurs in the general areas or switching on/off processes in energy, wide networks are performed. The strikes could induce voltages ranging from 500 to 5,000V and having a duration stretching from 1 ms to 10 $\mu$s, respectively.

According to IEC 60947-5-2, the inductive sensors must be capable of withstanding up to 1.2 kV (500Ω) of IEC 61000-4-5 surge. Hence, all pairs of sensor connections: +$V_B$, OUT, and GND should be capable of withstanding up to 1.2 kV of 1.2/50- to 8/20-$\mu$s surge pulses (Figure 11.9) with both polarities and provided by a surge generator with an impedance of 500Ω (Figure 2.11).

The test signal to certify the sensor surge immunity is a single high-energy pulse, which is a dual exponential one: the pulse has an exponential rise time to the peak value of 1.2 $\mu$s, followed by an exponential decay back to zero over a longer period of time. It has a duration of 50 $\mu$s in the pulse middle.

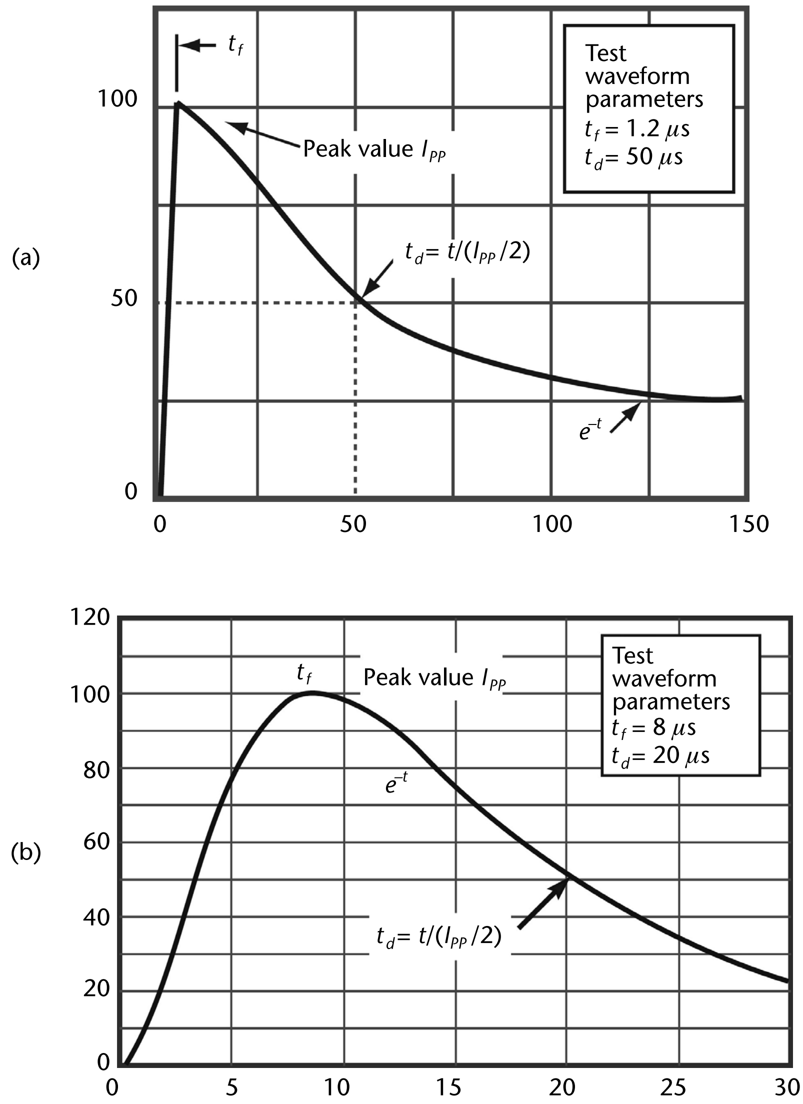


**Figure 11.9** Shape of surge test pulses: 1.2/50 and 8/20 $\mu$s, respectively [11].

This type test is performed with production release samples applying the pulse with both polarities between every two terminals. Additional conditions of the test (number of pulses, recovery times, etc.) are given in Section 2.2.4.

After every test step, the sensor is normally supplied and the normal operation test is performed.

To fulfill this hard energetic requirement, protection clamping diodes have to be installed in the sensor.

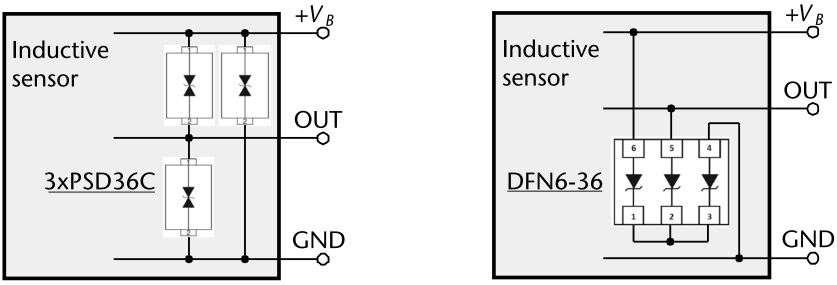
The popular procedure is to assemble suitable transient voltage suppressor (TVS) diodes connected between every pair of sensor terminals. They clamp the applied surge pulse on values below 60V, which can be absorbed by the electronics. The semiconductor industry is engaged to offer more and more compact packages with multiple TVS content. Figure 11.10 illustrates two elegant methods to connect suppressor diode arrays.

Note that a series of two diode structures is present for every pair of terminals and for both polarities. Each time one diode is forward-biased while the other is reversed-biased. Concluding, the clamping is bidirectional and a reverse polarity protection is fulfilled.

This procedure requires more space in the sensor housing and the parts are costly and sometimes incompatible with data communication.

A new trend, observed by up-to-date ASICs for sensors, is to integrate a protection structure on the sensor single chip or at least on the back-end chip. To save chip area, the star topology with floating central connection (CC) comes into consideration (Figure 11.11).

The integration of a clamping structure for high-energetic, short-time pulses (current peak up to 2A) on a silicon chip—and above all—beside sensitive sensor front-end electronics is a high design skill. Of course, the structure is exceptional in that it acts only when a high-energetic induced pulse appears on the sensor lines; otherwise, it is latent without any influences on the chip. Seen in this way, the design job seems to be easier as the design of an on-board integrated output stage (see Section 10.2.3). Nevertheless, the experience shows that high hurdles can be cleared successfully.



(a) (b)

**Figure 11.10** Surge protection TVS arrays: (a) three PSD36C units with maximum clamping voltage of 60V at 1A [12], (b) DFN6-36 device with maximum clamping voltage of 45V at 2A [11] and floating connection.

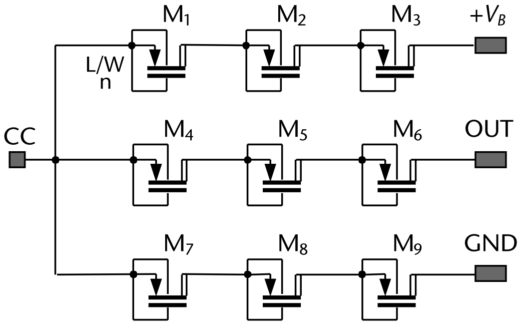


**Figure 11.11** Schematic of an integrateable surge-protection network (the protection units $M_x$, $x = 1 \ldots 9$, consists of $n = 110$ elementary P-channel MOS-transistors with $L = 0.7$ $\mu$m and $W =$ 2,500 $\mu$m).

The PSpice simulation tools are able to compute the electrical behavior of the structure during the surge pulse. However, local thermal effects (self-heating) and material-stressful occurrences during a surge-discharge are difficult to evaluate. In conclusion, the creation of a surge protection structure is a combination method consisting of computer-assisted design and experimental trail-and-error activity.

The bottom-up design of the Figure 11.11 circuit consists of four phases:

1. The protective mechanism is based on the voltage clamping by the breakdown voltage of a silicon junction of a transistor. Hence, the first step is to define suitable elementary transistor structures with varying parameters as well as to simulate and examine them.
2. Design of protection units containing more elementary transistors ($M_x$) that are connected in parallel. Their characteristics are number of elementary transistors ($n$), channel length ($L$), channel width ($W$), gate to source distance, etc. The most important feature is the resulting clamping voltage under surge current pulses (≈12V, usually).
3. Assembly of the best protection unit to get protection networks (e.g., Figure 11.11), layout design, mask manufacturing, test integration, and sample preparation.
4. Sample verification and structure validation performing surge tests in an EMC testing laboratory (see Section 2.2.4).

Note also that the measured effective clamping voltages are higher than the result of multiplication: the number of protection units times their clamping voltage. The drop voltages of the forward-biased units as well as the voltage drops on the metal conductive strips that cannot be neglected at currents of a few amperes have to be added to this amount. The increase of the clamping voltage lies in the range of a few volts.

The consumption of the chip area (1 to 2 mm$^2$) could be inadvertently considered to be a flaw (Figure 11.12). Nevertheless this circuitry needs definitely less room than TVS parts and has no additional connections lines on the sensor PCB. The chip area increase is less expensive than the external TVS devices.

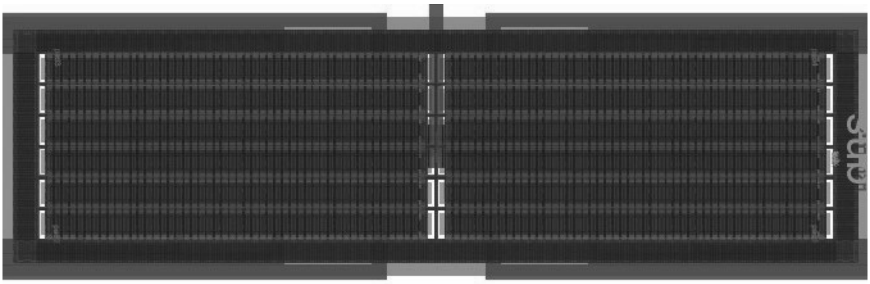


**Figure 11.12** Layout of an integrated surge protection structure.

## References

[1] Whitaker, J. C. (editor-in-chief), *The Electronics Handbook*, Second Edition, Boca Raton, FL: CRC Press/IEEE Press, 2005.

[2] http://www.ti.com/lit/ds/symlink/lm741.pdf, *Data sheet: LM741 Operational Amplifier, Texas Instruments*, 2015.

[3] http://www.ti.com/lit/ds/symlink/lm723.pdf, data sheet, LM723/LM723C, Voltage Regulator, Texas Instruments, 2013.

[4] http://www.ti.com/lit/ds/symlink/lm7800.pdf, data sheet, LM340, LM340A and LM7805 Family Wide VIN, 1.5-A Fixed Voltage Regulators, Texas Instruments, 2016.

[5] http://www.ti.com/lit/ds/symlink/lm317.pdf, data sheet, LM317, 3-Terminal Adjustable Regulator, Texas Instruments, 2016.

[6] http://www.onsemi.com/pub/Collateral/MC7900-D.PDF, data sheet, MC7900 Series, 1.0 A Negative Voltage Regulators, ON Semiconductor, 2013.

[7] Day, M., Understanding Low Drop Out (LDO) Regulators, TI-white paper SLUP239, 2006, http://www.ti.com/lit/ml/slup239/slup239.pdf, Texas Instruments.

[8] http://www.ti.com/lit/ds/symlink/tps769.pdf, data sheet, Ultralow-Power 100-mA Low-Dropout Linear Regulators, Texas Instruments, 2001.

[9] http://www.ti.com/lit/ds/symlink/tlv431.pdf, *TLV431x*, Low-Voltage Adjustable Precision Shunt Regulator, Texas Instruments, May 2018.

[10] Hering, E., and G. Schönfelder, *Sensoren in Wissenschaft und Technik–Funktionsweise und Einsatzgebiete*, Wiesbaden, Germany: Vieweg+Teubner Verlag / Springer Fachmedien GmbH, 2012.

[11] http://www.protekdevices.com/xyz/documents/datasheets/dfn6_36.pdf, data sheet, DFN6-36, 300 Watt Multi-Line TVS Array, ProTek Devices, 2015.

[12] http://www.protekdevices.com/xyz/documents/datasheets/psd.pdf, data sheet, PSD03–PSD36C, High Powered TVS Array, ProTek Devices, 2013.

## CHAPTER 12

# Inductive Sensors: Adjustment and Calibration

As already mentioned, the large majority of inductive sensors needs a final calibration—and sometimes—supplementary adjustments during the normal operation. This chapter provides an evolutionary presentation of the sensor calibration starting with the classical methods and ending with up-to-date procedures.

Typically, the calibration is done by modifying (trimming) a resistance element. For inductive sensors, this element is represented by the stage local trimming (LTM) in the generic sensor diagram (Figure 7.1).

The adjustable stages of inductive sensors (oscillator, comparator, etc.) are designed to make possible a functional trimming. By means of a resistive element, the trimming occurs with the sensor in normal operating mode and monitoring its output signal (teach-in) instead of the resistance value. This resistor is consistently present in all discussed oscillators in Chapter 8 for inductive sensors (e.g., the resistor $R_{ADJ}$ in Figure 8.34).

The sensor industry relies on a variety of resistance trimming techniques. The present toolbox includes laser trimming resistors, manual mechanical potentiometers, fusible passive resistor arrays, and electrically configurable resistor arrays (digital potentiometers and other methods involving nonvolatile memories). In practice, there is no perfect single solution.

For presenting the various adjustment procedures of the inductive sensors, we propose an innovative rating, where the main criterion is partially unusual and is the required remoteness (Table 12.1). The larger this access way (mechanical, electrical, etc.), the more suitable the method for inductive sensors.

In addition, the trimming of inductive sensors is critical. Using the traditional methods listed above, the trimming occurs very soon in the fabrication line, in an open state of the sensor, when a direct access to the trimming element on the PCB is possible. To provide a high immunity against various disturbing environmental sources (Figure 2.7), the sensor is hermetically sealed and encapsulated in an enclosure (usually metallic). These components and fabrication steps—after already being trimmed—sometimes dramatically influence the internal sensing element and lead to misalignment, total decalibration, or malfunction of the sensor. The lower the sensor functionality headroom, the higher these negative effects.

The conclusion is the need for trimming procedures, which can be performed when the sensor is completely assembled. Their presentation is the main topic of this chapter.

**Table 12.1** Benchmarking Conventional Trimming Procedures

| *No.* | *Means* | *Procedure* | *Degree of Accessibility* |
|---|---|---|---|
| 1 | Trimmable resistor | Resistor trimming by laser beam | Very low; the method requires mechanical contact to the resistor |
| 2 | Manual mechanical potentiometer | Manual adjustment by moving of the potentiometer wiper | Very low; the method requires direct mechanical link to the potentiometer |
| 3 | Rejustor™ | Electrical in-circuit adjustment of the Rejustor resistivity by thermal processing | Low; the method requires an electrical contact to the Rejustor heating terminals |
| 4 | Digital potentiometer | Electronical trimming by transfer of a binary code via a serial interface | Medium; the method requires an electrical data link (2 to 4 wires) to the device interface |

## 12.1 Traditional Sensor Trimming Procedures with Commercial Components

Despite the disadvantages specified above, the traditional procedures still capture the sensor fabrication lines due to their lower production costs.

### 12.1.1 Trimmable Resistors

The resistance change of a trimmable resistor by invasive structure modification—in the past using sandblast and recently perfected with laser beam—is the most applied trimming procedure in serial production of mass-produced devices (e.g., inductive proximity switches). The technique enables parametric values (e.g., effective operating distance $s_r$, see Section 2.1.1.2) to be achieved by trimming the resistance value. This substitution for trimming potentiometers or manually selected fixed resistors is by far more reliable, faster, and cheaper.

The achieved benefits are high reliability without movable parts, pick-and-place capability, cost reductions, PCB space savings, and full automatic inline trimming process with reduced trimming time.

A trimmable resistor is basically an untrimmed thick-film chip resistor consisting of a resistive layer, which is printed on a ceramic substrate and is followed by an overglaze (Figure 12.1(a)). The lateral terminations contain a precontact and a nickel barrier covered by a pure tin finish.

The resistance can be increased if a trimming kerf is cut in the resistive layer and the material surplus is vaporized with a controlled laser beam.

If the monitored sensor output is a switching output, this signal can control the laser directly. The functional trim runs without interruption and is stopped if the output reacts, when the propagation times in the sensor and laser are short enough to get a result without overdrive delay errors. Otherwise the adjustment will be made step-by-step, performing a loop activity: trim, measure, compare achieved and wanted sensor parameters, and then decide whether to do a new iteration or not.

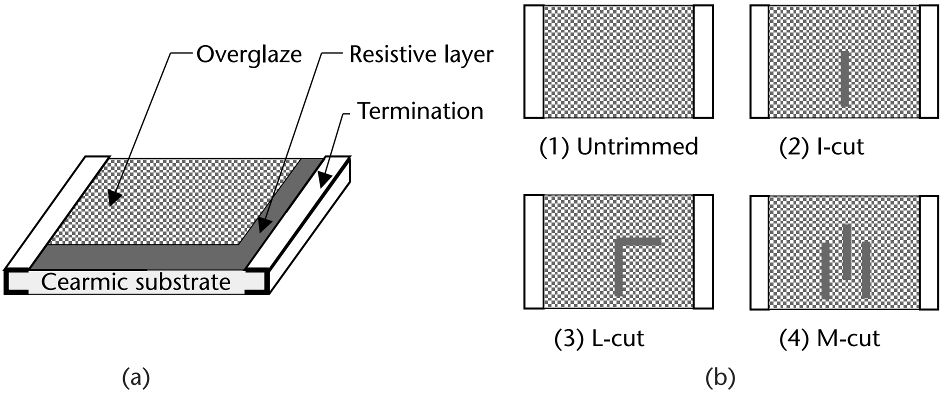


**Figure 12.1** (a) Typical construction of a trimmable resistor, and (b) popular trimming cuts (top view) [1].

Many manufacturers provide a wide variety of resistors with different performances. Reference [1] gives a good summary of these features:

- Resistor sizes: 0603, 0805, or 1206 (0603 has, for example, dimensions of 0.06 × 0.03 inches);
- Resistance range $1\Omega$ to $10\ \mathrm{M}\Omega$;
- Tolerances: ±5%, ±10%, ±20%, etc.;
- Rated power: $P_{70} \leq 250$ mW (dependant on chip size);
- Temperature coefficient: 50 ppm/°C to 250 ppm/°C;
- Endurance (1,000 hours at $P_{70}$): 1% to 2%.

The most important feature is the trim factor, which is the ration of the trimmed value to the initial untrimmed value and should be as great as possible.

The cut shape and geometry (Figure 12.1(b)) strongly influence the trim factor, trimming accuracy, long-term stability, and reliability. I-cuts are the simplest and fastest; other than that, they offer modest features and a trim factor below two. The L-cut seems to be the best compromise solution with good speed, high accuracy and stability, and an acceptable factor of two. Finally, the complex M-cut can be applied to achieve trim factors up to 20 with acceptable parameters.

The reader can find in [1] additional information and concrete data regarding the trimming equipment, laser setting, and trim process flow.

In the case of ceramic thick-film hybrid circuits, the resistor could be directly printed on the ceramic substrate and trimmed like mentioned above. Due to the large tolerances of the printed resistors, they need to be preadjusted before hybrid delivery, which means additional costs.

From the author's experience, the resistor trimming method has some drawbacks. Apart from the fact that this invasive procedure could be polluting and harmful, it provides only one-time adjustments that are irreversible and without a fine-tuning opportunity. The biggest drawback is the need for direct access to the trimming resistor. For inductive sensors this is a fundamental disadvantage. The trimming has to be made before the sensor is assembled and all negative influences of the following steps in the final assembly or of metallic constructive parts are not compensated.

The benchmarking to determine if the trimming resistor offers a better solution in comparison to electronic methods in the sensor manufacturing has been an open question for more than 20 years. Every party brings solid pro arguments; meanwhile the adjustments of general electronic devices have become exclusively electronic. At the end of this chapter, the reader will have the possibility to get involved in this constructive opinion dispute.

### 12.1.2 Rejustor

To prevent the need for a direct contact to the resistive layer of the trim resistor, Microbridge® Technologies Inc. proposes to replace the mechanical laser intervention by a thermoelectrically induced change of the material resistivity [2].

The Rejustor™ (electronically readjustable resistor) is a passive, MEMS-compatible, adjustable microresistor. It is nonvolatile, meaning that it does not need power to hold its adjustment. It is readjustable many times (thousands of repeated trimming operations) to very high precision in a few seconds (e.g., 0.1% to 0.002%, depending on a variety of factors).

It uses only electrical signals to thermally modify the crystal structure of polysilicon resistors, thereby changing their resistance. Once changed, the crystal structure remains intact until deliberately heated and changed again.

Rejustors can be adjusted in a range at least 30% down from their nominal manufactured value.

The Rejustor technology is based on standard CMOS chip technology (Section 7.3.2) with simple postprocess cavity-etching, which releases the suspended microstructure (Figure 12.2)(a)). These offer enhanced thermal isolation from the substrate as well as from the low mass. This enables localized, controllable, and rapid thermal cycling of electrical elements embedded in the microstructures.

One or two trimmable functional resistors ($R_1$ and $R_2$ in Figure 12.2(b)) with the associated secondary adjustment resistors are realized on the microstructure. These are used as heater, and are powered during the actual trimming step. For this purpose, their terminals R1ADJ and R2ADJ and the heating ground HGND will be connected to a Matchbox [3] that coordinates the trimming manually or computer added.

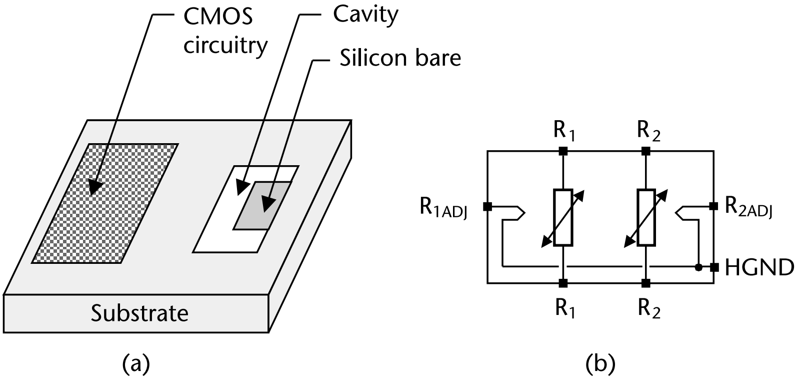


**Figure 12.2** (a) Schematic construction of a Rejustor, and (b) circuitry of a dual Rejustor package [2].

The adjustment of the functional resistances is accomplished by locally heating the auxiliary resistors, gently changing the physical properties of the functional resistors with each heat cycle. The algorithms used to apply heat cycles are adaptive, involving a repeating sequence of measuring the functional resistance at room temperature, computing pulse parameters for the next cycle, and applying of the high-temperature sequence.

By using only electrical signals, Rejustors can also be used in-circuit to perform trimmings. Adjustments can be carried out at low voltage and low current before and/or after packaging.

Rejustors can be temperature-coefficient matched with other Rejustors. Using standard CMOS polysilicon, they can be trimmed from ±100 ppm/°C down to within 2 ppm/°C or deliberately mismatched (e.g., by ±50 ppm/°C, for balancing and temperature conditioning).

Features of the Rejustor include

- Packages: 8-lead SOIC or 16-pin QFN (3 × 3 mm);
- Wide range of initial R values (dual values in the package: 2 × 4.7 kΩ to 2 × 90 kΩ, or 10 + 20 kΩ to 10 + 90 kΩ);
- Long-term stability;
- Frequency range: DC to 1.5 GHz;
- Temperature range: −55°C to +125°C;
- No external memory or boot-up is required;
- No moving parts; thus Rejustor is vibration insensitive.

An enhancement to the technology is the eTC Rejustor from Microbridge, which offers a solution for all-passive, all-analog, electrically controlled, temperature-coefficient adjustment for postpackage temperature conditioning. With an eTC Rejustor, each resistor's resistance and temperature coefficient of resistance TCR can be adjusted to independent targets.

This allows flexibility for controlling temperature-related problems. Amplifier offset and TC offset, for example, can be compensated in the analog domain with continuous value adjustment. The adjustment is done electronically after board assembly, and thus the design engineer can wait for all other parameter variances and temperature sensitivities to manifest and then simply null out the cumulative effects during the final test.

### 12.1.3 Manual Mechanical Potentiometers

A potentiometer is an electrical resistor whose resistance can be mechanically varied. It has a minimum of three terminals and can be used either in potentiometer mode (voltage divider) or in *rheostat mode* (variable resistor with only two terminals). The physicist Johann Christian Poggendorff invented the potentiometer.

The potentiometer contains a supporting dielectric core piece, which carries a resistive layer with two connection terminals at its ends and a movable sliding contact (the wiper). The wiper performs a mechanical division of the electrical total resistance in two partial resistances whose sum is permanently constant und equal to the total resistance.

Potentiometers are available as rotary or slide potentiometer for general adjustment tasks.

Three conventional resistive layers are established:

1. *Wire wound*. A resistance wire is wound around an insulated core. This is then fixed in place and bent to the design shape. Precision manufacturing methods allow reliable processing of resistance wires down to an extremely small diameter of 0.01 mm. A wiper contact slides on the resistance wire surface and picks up the voltage signal at the relevant point. Due to the operating principle, the resolution is dependent on the number of wire windings.
2. *Conductive plastic*. Using screen-printing methods, a thick-coat resistance track is applied to various carrier materials (FR4, ceramic, plastic, etc.). The coating is then hardened in an oven. The wiper slides on the resistance track and picks up the relevant voltage signal. Due to its operating principle, the resolution is infinite. However, conductive plastic potentiometers are not suitable for all applications.
3. *Cermet conductive substrate*. A cermet is a composite material consisting of ceramic (cer) and metallic (met) materials. Using a screen-printing process, a thick coating of metal/metal oxide is printed onto a ceramic carrier material. This layer is hardened at high temperature. Otherwise, cermet potentiometers are identical to the plastic ones. Their resolution is also infinite.

For all realizations, the resistive element (resistance wire, thick-coat resistance, or cermet resistance) and the corresponding wiper system are accurately matched to ensure maximum reliability throughout the entire lifetime of the system. Manufacturers recommend wire wound, and cermet versions are primarily suitable for general use.

Features of the potentiometers refer to

- The dependence function between resistance and wiper angular or linear position: linear or nonlinear (positive or negative logarithmic);
- The nominal value between the main terminals (end-to-end resistance) and its tolerance;
- The linearity and repeatability. High-quality linear potentiometers offer a linearity of up to ±0.15% and repeatability of a maximum of 0.1°;
- The maximum current density in the conductive layer, the voltage strength and rated power;
- Because of wear of the resistive material caused by the wiper that is related to their operating principle, potentiometers are not wear-free. Hence, the lifetime (number of cycles in the range of a few millions), endurance, and number of wiper movements are relevant.

For functional trimming of sensors a particular potentiometer design, namely the trimming potentiometer (trimmer), is used. This is structured for a small number of actuations. After adjustment, the wiper position is secured (e.g., by a drop of paint).

As the with a standard potentiometer, trimming potentiometers are manually adjustable, variable resistors with three terminals and different number of turns.

Two terminals are connected to a resistive element and the third terminal is connected to the movable wiper. They can also be used as voltage divider or as rheostat.

Multiturn trimmers consist of three actuation types: lead screw, worm gear, and planetary gear. Lead screw actuated trimmers have a linear wiper path. Worm gear and planetary gear actuated trimmers have a circular path. Multiturn trimmers (4 to 25 turns/200 cycles, typ.) are suitable for applications that require fine resistance adjustment.

Single-turn trimmers are actuated by a direct-drive mechanism. It consists of a simple rotor with a slot, whose wiper movement is traversed in a circular path. Package types available include (1) sealed trimmers, which are built with an O-ring seal that prevents entrance of moisture and serves as a mechanical restraint to prevent unwanted wiper movements, and (2) open-frame trimmers, which are not sealed to withstand a cleaning process and are usually lower in cost. Single-turn trimmers are designed with a direct-drive actuation for resistance adjustment, where the level of accuracy is not critical.

Trimming potentiometers are available in miniaturized sizes as

- 3/4-inch, rectangular body, multiturn, for mounting types: through-holes or SMD;
- 4-mm, square body, multiturn cermet, SMD;
- 3/8-inch square body, multiturn, through-holes;
- 3/4-inch, cylinder body, 4-turn, through-holes;
- Round or cubical, single-turn for through-holes or SMD mounting.

The manufacturers are engaged to permanently improve the trimmer features. Trimpot® series [4] provides high-level features such as

- Independent nonlinearity (for definition see Figure 1.4 (a)) with an error ≤1.5% (standard trimmers usually 4%, typically).
- Contact resistance variation (CRV) ≤0.2% instead of 0.8%. CRV gives the maximum instantaneous change of the contact resistance if the wiper position is changed.
- Theoretically unlimited low resolution. Practically, this is defined by the low CRV.

Over and above, a third type, the high-quality calibration potentiometers could be used, if needed, to perform a metrological sensor adjustment. First, the potentiometer is connected to the PCB pads for the adjustment resistor and the sensor is adjusted. The potentiometer is removed and the calibration value is read out. Then, a fixed resistor having exactly this value is pulled out from an assortment of fixed resistors and soldered on the PCB.

### 12.1.4 Digital Potentiometers

Digital potentiometers are digitally controlled devices that provide the same analog functions as a mechanical potentiometer.

They offer more accurate, more robust, and faster sensor functional trimming with smaller voltage glitches. There is a large variety of these devices including types with different memory technologies, single- or dual-supply, different digital interfaces, high- and medium-resolution, and the industry's broadest end-to-end resistance.

Apart from the digital and supply terminals, the digital potentiometer is a three-terminal device (Figure 12.3(a)), with an internal architecture that contains an array of resistors and switches (FET transistors). The resistors are connected in series between the end terminals A and B. The wiper terminal W is digitally programmable to access any of the $2^n$ tap points on the resistor string.

The end-to-end resistance $R_{AB}$ between terminals A and B usually offers spanning from 1 kΩ to 1 MΩ. The partial resistances between one end terminal and wiper, namely $R_{AW}$ and $R_{WB}$, are complementary and progress in opposite directions (increase/decrease).

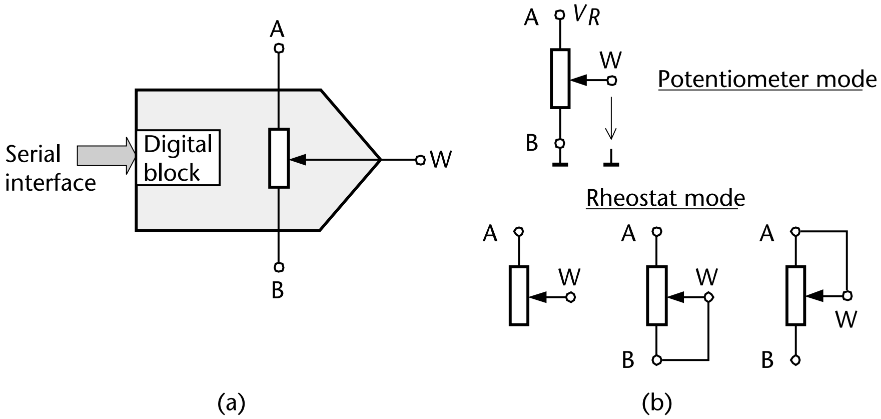
The resistive part of the potentiometer can be considered to be isolated from the control electronics. There is no restriction on the voltage polarity applied to the terminals A, B, or W or in the current flow direction. The limitations refer to the applied voltage, which does not exceed the power supply rails and to the resulting current density, which should remain below the specified maximal value (on the order of a few mA).

Digital potentiometers can be used as a potentiometer or as a rheostat (Figure 12.3(b)).

In the potentiometer mode, the device is supplied with a reference voltage $V_R$ and acts as a voltage divider. The wiper voltage is proportional to the reference and can be calculated as:

$$V_{W\_GND} = \frac{\text{CODE}}{2^n} \cdot V_R \tag{12.1}$$

where $n$ is the potentiometer resolution (see Table 12.2), $2^n$ is the taps number, and CODE is the value of the digital code transferred to the potentiometer via digital interface.



(a) (b)

**Figure 12.3** (a) Schematic block diagram of a digital three-terminal potentiometer, and (b) connection options for potentiometer or rheostat.

In the rheostat mode, the device operates a variable resistor between the wiper and an arbitrary end terminal. The unused end terminal can be left floating or tied to the wiper terminal (Figure 12.3(b)). The rheostat resistance can be calculated by:

$$R_{AW} = \frac{2^n - \mathrm{CODE}}{2^n} \cdot R_{AB} + R_W \text{ or } R_{WB} = \frac{\mathrm{CODE}}{2^n} \cdot R_{AB} + R_W \qquad (12.2)$$

where $R_W$ represents the wiper's first connection at the B terminal for zero scale and has a typical value of $R_W = 70\Omega$ for the family digiPOT [5].

Digital potentiometers support more digital interfaces: standard serial as SPI or $I^2C$ (see Section 14.5) but also push-button (two low-active, up/down inputs).

Four types of integrated memories are implemented in digital potentiometers to control the wiper position. Additionally, they allow the user to set the wiper's power-on reset POR position. The position can be reprogramed but always returns to the POR-position on power-up.

The memories provide the following functionalities:

- Volatile memory typically causes a power-up to midscale.
- One-time programmable (OTP) memory allows the user to program the wiper power-up position once. This microfuse using memory is a low-cost execution but nonetheless an application-critical option. A first faulty programming makes the device unusable and the sensor nonadjustable;
- Multitime programmable (MTP) memory is a cost-effective option, which supports 2 to 50 wiper programming cycles.
- The integration of an EEPROM is obviously the best option but is also the most expensive. The up-to-date technologies guarantee attractive features: up to 1,000,000 programming cycles and data retention of 50 years at 125°C [5]!

The resolution of the resistance values (see Section 1.3.4) depends on the specific end-to-end resistance ($R_{AB} = 1\ \mathrm{k}\Omega$ to $1\ \mathrm{M}\Omega$) and on the number of converting bits ($n = 5$ to $10$). Theoretically, the LSB step size is expressed by:

$$STS_{LSB} = \frac{R_{AB}}{2^n} \qquad (12.3)$$

A remarkably wide variety of integrated digital potentiometers (more than 50 types) is singly offered by Analog Devices [5]. As an example, Table 12.2 summarizes and compares two telling circuits (both single-channel).

The manufacturer openly states the inherently large error of the absolute end-to-end resistance (±20%, typically), which could be an impediment for some applications, and makes improvement proposals. Some devices contain their individual error stored by the manufacturer in an additional memory location.

In addition, the resistance combined with the parasitic capacitances of the switch, pins, and board creates an RC low-pass filter, which determines the maximum AC frequency that can be passed through the potentiometer before it is attenuated by more than −3dB. A device with a lower resistance $R_{AB}$ can support a higher bandwidth: 5 MHz at 1 kΩ, 1 MHz at 10 kΩ, but only 6 kHz at 1 MΩ [5].

**Table 12.2** Benchmarking of Two Representative Single-Channel Digital Potentiometers: digiPOT

| *Features* | *AD5291* | *AD5259* |
|---|---|---|
| Resolution ($n$) | $n = 8$ (256 positions) | $n = 8$ (256 positions) |
| End-to-end resistance $R_{AB}$ (kΩ) | 20, 50, 100 | 5, 10, 50, 100 |
| Specific $R_{AB}$ tolerance stored in EEPROM | n.a. | Yes (0.1% accuracy) |
| Temperature coefficient (ppm/°C) (potentiometer/rheostat mode) | 5/35 | ≤300 |
| Digital interface (see Section 14.5) | SPI | $I^2C$ |
| Memory type, programming cycles | MTP, 20 (internal fuses) | EEPROM |
| Typical data retention at 55°C | n.a. | 100 years |
| Single-supply operation (V) | +9 to +33 | +3 to +5 |
| Dual-supply operation (V) | ±9 to ±16.5 | n.a. |
| Operating temperature (°C) | −40°C to +105°C | −40°C to +125°C |
| Package | 14-lead TSSOP | 10-lead LFCSP, MSOP |

*From:* [6, 7].

This narrow bandwidth could be an impediment for implementations in inductive sensors, which work at high frequencies.

## 12.2 Specific Programmable Electronic Devices Used to Calibrate Inductive Sensors

The classical methods described in the previous section require direct view between the laser beam and the trimming resistor or mechanical contact to wiper or electrical connection lines. Digital potentiometers could be overdesigned for serial products, need additional supply power, and increase the sensor price.

The best alternative, confirmed by the high number of examples, is to design and implement a just-enough integrated trimming device that totally and optimally fulfills the demands of inductive sensors.

The experience concerning inductive sensors, which is gained from the previous chapters, allows us to specify this device by following particular as well as essential features:

1. The ASIC has to perform a tech-in trimming of the inductive sensor, which is finally assembled in its metal protection housing and during its normal operation. The reason is to compensate all the parasitical influences on the sensing element, which generates parameter deviations.
2. This trimming should be possible at more locations [8]:
   a. In the production line for tests and default adjustment with the standard target;
   b. In the supplier warehouse to prepare precustomized sensors;

   c. By the user for final adjustment and fine tuning on the metallic object in the application. If during the sensor operating lifetime, the machine part in which the sensor is mounted suffers a mechanical misalignment, this can be compensated by performing a new sensor calibration. Hence, OTP versions are adverse. Calibration via an EEPROM, which is by default programmed during a sensor factory test, and can be readjusted in situ on demand, is practically the desirable demand.

3. To maintain the compliance with the standards, the sensor does not have additional connection terminals (connector pins, cable wires, etc.). The ideal solution would be to implement a reliable serial communication for the programming commands and to use a programming unit (temporarily connected to the sensor lines, only during teaching, to supply and program it). Smart solution which uses already existing data channels will be presented in this section.
4. The intelligent link assumes a communication protocol (see Section 14.1). This should guarantee high data immunity and have a strong limited instruction set that contains specific commands such as trimming resistance up/down, or max/min, among others. An irreversible lock (channel close) command is implied;
5. Equally important are the technical characteristics: a chip as small as possible, available for miniaturized assembling techniques, chip on board, flip chip, and so forth (see Section 7.3.1.2), cost-effective.
6. The resistive ASIC part should have a low temperature coefficient and a wide −3 dB bandwidth.
7. The ASIC should be a low-voltage, low-current consumption system that is also suitable for two-wire sensors. The data retention should be high, to prevent memory lost data, and hence, sensor misalignment in the application.
8. The calibration techniques define two classes: one, which uploads adjustment coefficients, and the other, which performs self-teaching adjustments, as the tech-in procedure.

### 12.2.1 Established Methods for the Trimming of Inductive Sensors

A good source of information regarding inductive sensor trimming solutions are the databases of the international patent classes G (physics) and H (electronics). As a result of a long-term, permanent patent publication surveillance, this section provides a short overview of some significant ideas.

These innovative ideas partially fulfill the requests listed above and represent a real progress. However, they are not the ideal solution to meet all demands.

Omron Corporation, Japan, created a circuit that could enhance the electronic of an inductive sensor, replacing the classical resistor $R_{ADJ}$ by a resistor network. The circuit contains (1) a 4-bit parallel resistive network $R_1 \ldots R_4$ with four switches $T_1 \ldots T_4$, (2) a series of four microfuses $F_1 \ldots F_4$ controlled by four switching transistors $T_5 \ldots T_8$, and (3) a digital part with a counter consisting of four flip-flops $FF_1 \ldots FF_4$ (Figure 12.4).

Before starting trimming, the standard target is placed in front of the sensing element at the distance to be programmed (teach-in procedure) and the counter is

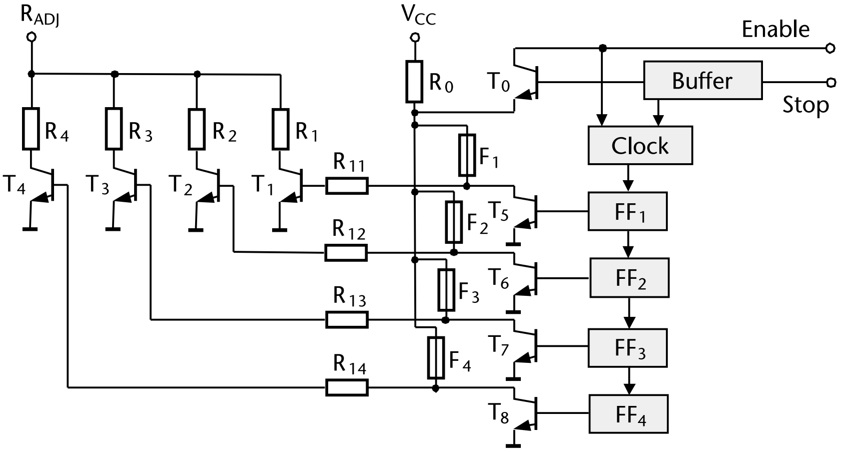


**Figure 12.4** Schematic block diagram of an OTP trimming circuit

cleared. Correspondingly, the network switches $T_1$ ... $T_4$ are saturated and $R_{ADJ}$ has the initial lowest trimming value, which can be expressed (neglecting the transistor resistances) by:

$$R_{ADJ} = \frac{1}{G_1 + G_2 + G_3 + G_4} \tag{12.4}$$

where the $G_i$ are the corresponding conductances of $R_i$ ($i = 1$ to 4) ($G_i = 1/R_i$).

The trimming occurs in three steps:

1. The process is initiated by pulling the signal ENABLE on HIGH. The clock generator starts and the counter begins to count up the clock pulses. The binary count-up of its flip-flops correspondingly saturates the transistors $T_5$ ... $T_8$. Every saturated transistor of this string ties down the base of the partner transistors $T_1$ ... $T_4$, which temporarily disconnect their collector resistor. Thus, $R_{ADJ}$ evolves on a digital increasing ramp. The main transistor $T_0$ is nonconductive. This activity is stopped when the mechanically set distance is sensed and the sensor electronics (e.g., a comparator) replies with the signal STOP.
2. The counter running is stopped and $T_5$ ... $T_8$ store the digital trimming value. Then, $T_0$ is activated via buffer, and carries high currents through these fuses, which are grounded via saturated transistors $T_5$ ... $T_8$. These fuses blow and the corresponding bases of $T_1$ ... $T_4$ are definitively separated to $V_{CC}$.
3. Finally, the command signals are removed, and the transistors $T_5$ ... $T_8$ lose the control function. The final trimmed network state, and thus, the value of $R_{ADJ}$, is determined by the transistors $T_1$ ... $T_4$ that remain supplied through the nonmelted fuses.

The schematic is simple; it can be realized in discrete or integrated form, allows a high speed trimming, and has great electromagnetic immunity.

At the same time, it has considerable drawbacks: the OTP character and the need for additional external lines (ENABLE), the nonlinear evolution of the $R_{ADJ}$ ramp due to the parallel topology, an infinite end-value equivalent to an interruption of the $R_{ADJ}$ path, and a high temperature dependence of $R_{ADJ}$ due to the saturated transistors in series with the real resistors.

As is well known, integrated resistors need large silicon areas. In order to keep the chip area in limits, the nucleus of the circuit is switched over to the digital functionality, which needs less area. The result is a dynamic trimming method instead of a static one. The method is based on a pulsewidth modulation (PWM) with a variable degree of modulation and requires an internal signal detection with integrating RC filter [9].

The trimming network is now reduced to a series circuit consisting of a fixed preresistor $R_1$ and a second resistor $R_2$, which is periodically short-circuited by the FET switch $T_{SW}$ (Figure 12.5). The result of this PWM is an average resistance value expressed by:

$$\bar{R}_{ADJ} = \frac{1}{T}\int_0^T R(t)\,dt \approx R_1 + R_2(1 - D) \tag{12.5}$$

where $D$ is the degree of modulation, which controls the trimming value and is expressed by:

$$D = \frac{t_{ON}}{T} \tag{12.6}$$

with $t_{ON}$ the switch on time of $T_{SW}$ and $T$ the repeat period. To conclude, the trimming of a wanted effective operating distance can be done by a suitable adjustment of $D$.

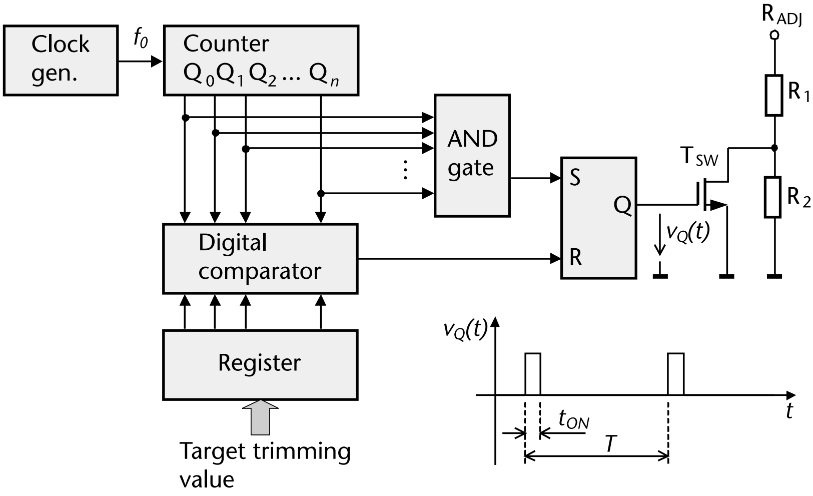


**Figure 12.5** Schematic block diagram of a dynamic trimming circuit with PWM.

The switch $T_{SW}$ is toggled with a pulse signal hawing the duration $t_{ON}$ and the period *T*, which is supplied by the output *Q* of an RS flip-flop.

The flip-flop is periodically set if the counter that counts clock pulses with the frequency $f_0$ becomes full.

The reset point of time is variable. The periodical reset occurs if the digital comparator signalizes a coincidence between the digital content of a preset register, which defines the target trimming value, and the instantaneous counter digital value.

The counter bit number *n* defines the resolution and hence, the repeat period according to:

$$T = \frac{2^n}{f_0} \tag{12.7}$$

The higher the bit number, the better the resolution but the lower the system dynamic.

Despite the complexity of the digital part, the circuit can be integrated using MOS processes (Section 7.3.2) on a small chip area. The clock frequency $f_0$ should be high enough in comparison to the sensor switching frequency. The trimming characteristic; that is, the resulting operating distance versus digital preset value is nonlinear due to ripple of the $\overline{R}_{ADJ}$ This disadvantage can be minimized by choosing $R_1 >> R_2$. However, this measure reduces the trimming range.

It is interesting to exemplify how inventive the companies are in order to find innovative solutions and then to protect them by granted patents. One competition field was to have a protected proprietary solution to write commands into the sensor and eventually to read confirmation data.

The manufacturer IFM electronic GmbH/ Germany created a general digital trimming structure consisting of a $R_{ADJ}$ resistor network with switches. The switches command is provided in normal operation by a nonvolatile memory and during the trimming by a counter.

The counter has two internal inputs: count-up or -down. To avoid additional lines, IFM proposes to act on these inputs via external light signals → sensor LED optical indicators → internal interface between LEDs and inputs. Typically, LEDs are light-emitting devices and are used as sensor optical indicators. However, they provide a low-level photovoltaic effect in no supplied status, which can be used to write commands in the sensor. The method has limits; the programming has to be correlated with sensor states in which the LEDs are not activated. To prevent false programming, the communication immunity has to be assured by suitable coding, modulation, and spectral filtering measures.

Siemens AG/Germany has an alternative proposal. A similar counter has again two inputs: one for normal operation/programming, the other for counter-start/-stop. To generate these signals, the inventors propose to bring the sensor into normal operation and to artificially provoke a short circuit at the sensor short-circuit protected output. This action can be coded internally as a result of a coded instruction recognition.

It is interesting to note that Siemens is one of the initiators of the IO-Link communication system (Section 14.7), and the short-circuit procedure is also implemented here for the wake-up event (Figure 14.13).

An electromagnetic coupling between an external transmitter and an internal receiver (at the limit, the LC circuit) is also an idea, as long as the inductive sensor has a non-ferromagnetic housing or its operation is not influenced by the programming electromagnetic field.

The overview of the electronic trimming methods may be completed noticing how important the trimming schematic with microcontroller architectures could become for several classes of inductive sensors. In spite of the plethora of hardware solutions, the microcontroller takes over several functions for the software stored in its memories. In general, the domain is highly ongoing and professional publications are much spread. The block diagram in Figure 12.6 tries to represent the common denominator of these innovative proposals, emphasizing the following characteristics:

- After the primary signal processing performed by the inductive sensing element (eventually supported by a minimal analog electronics), the signal path passes through the microcontroller up to the final power amplification in the output stage (similar to Figure 7.29).
- Particular functions, such as the linearization, temperature compensation (Chapter 13), are also carried out by the microcontroller.
- Regarding programming tasks, the microcontroller can perform the classical trimming. In addition, it can realize the division of the measuring range into more zones, hysteresis changes, or prevention functions indicating subranges with low headroom/safety margin. Several forms of diagnosis can be implemented [10].

Notwithstanding this trend, the specific solutions with ASICs are the market leaders. It is important to underline that the company's activity in the field of new designs and methods for inductive sensor trimming is still in progress today, despite the dominance of microcontrollers [11, 12].

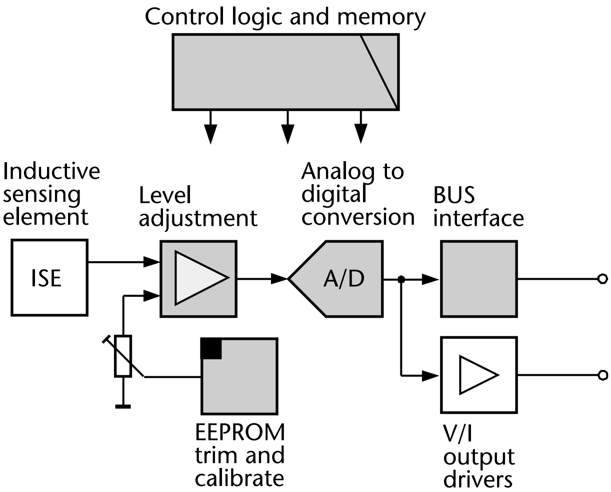

Control logic and memory
Inductive sensing element
Level adjustment
Analog to digital conversion
BUS interface
ISE
A/D
EEPROM trim and calibrate
V/I output drivers

**Figure 12.6** A microcontroller architecture (gray shaded stages) in an inductive sensor circuit. An EEPROM trims and calibrates the input threshold to match it to the signal from the sensing element (ISE). The ISE output signal is prepared by an ADC for transmission to a processing unit.

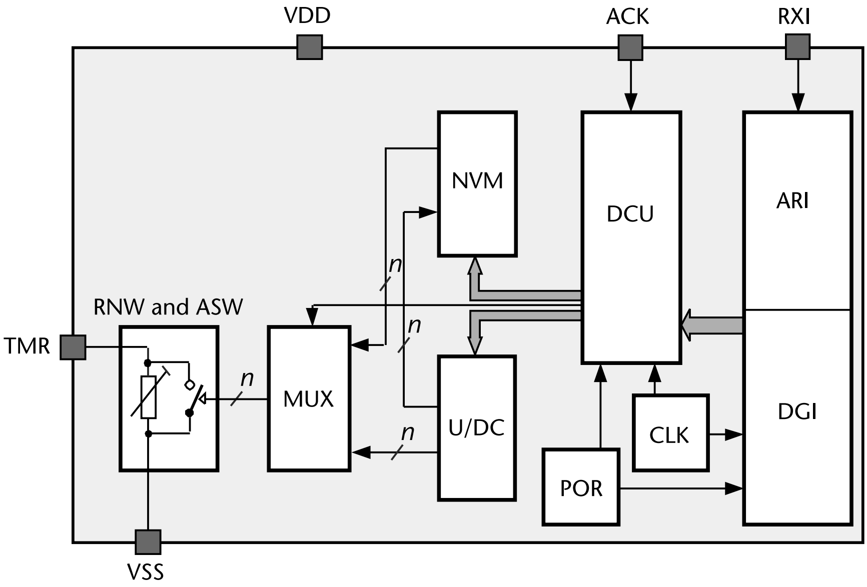
### 12.2.2 ASIC and ASIC Sections for Trimming of Inductive Sensors.

Consequently, following the presentation style to those in Section 7.1, the present section begins with a generic functional diagram of ASICs for the trimming of programmable inductive sensors (Figure 12.7). The structure should be modular and capable to perform more tasks—first, the sensor trimming (Chapter 12) and second, temperature compensations (Chapter 13).

The structure contains three main units:

1. Actuating unit (ACU) that is the interface of the trimming ASIC with the electronics of the sensor to be trimmed. This actuating part consists of the following two strings of passive and active electronic devices:
   a. Resistor network (RNW), which realizes the variable resistance;
   b. Analog switches (ASW), which affects the step-by-step adjustment of RNW size.

   Subsection 12.2.2.1 treats methods to implement and optimize the ACU.
2. Control unit and peripherals (CU/P) is the central ASIC unit and performs the specific ASIC functionality. On this part, the CU/P specifically consists of following electronic stages:
   a. Up/down counter (U/DC), to perform a rising or falling digital ramp, which finally controls the ASWs.



List of acronyms and initialisms:

**ACU: Actuating unit**
ASW: Analog switches
RNW: Resistor network

**CU/P: Control unit and peripherals**
U/DC: Up/down counter
NVM: Non-volatile memory
MUX: Digital multiplexor
DCU: Digital control unit
CLK: Clock generator
POR: Power-on reset

**C/PU: Communication and programming unit**
ARI: Analog receiving interface
DGI: Digital interface

**Figure 12.7** Generic functional diagram of ASICs for the trimming of programmable inductive sensors.

   b. Nonvolatile memory (NVM) in the form of an EEPROM or FRAM, which borrows at the end of a new trimming the resulting U/DC content and stores that for a sensor operation time period, theoretically infinite.
   c. Digital multiplexor (MUX), which acts as a path switch, connecting the U/DC with the ASW during the trimming process or the NVM to transfer the ASW control to the memory out of the trimming time (normal steady state).
   d. Digital control unit (DCU), which is the core piece and coordinates the entire functionality of ASIC. It should be a state-machine, as described in Section 12.2.2.2.
   e. Clock generator (CLK) and power-on reset (POR) constitute a sufficient minimal DCU periphery, which is also detailed in Section 12.2.2.2.

3. Communication and programming unit (C/PU) marks the low-signal and high-sensitivity part of the sensor. The C/PU consists of the following electronic stages:
   a. Analog receiving interface (ARI), which receives the analog signals that typify the programming commands and convert them into digital form (Section 12.2.2.4).
   b. Digital interface (DGI) borrows these serial digital frames and decodes them in command signals for the DCU.

An embodiment of this generic circuit is schematically represented in Figure 7.22 and its operation is presented in Section 7.4.2. The representation is reloaded in Figure 12.8 using the same initialisms as those in Figure 12.7. The device is indirectly supplied from the sensor master electronics via $V_{DD}$ and the programming input RXI is connected to the sensor supply terminal $+V_B$ via a coupling capacitor.

Also, the desired suitability for two-wire sensors imposes hard supply conditions: $V_{DD} = 2 \ldots 3$V and current consumption $I_{DD} \leq 50\ \mu$A, and restricts the availability of the integration processes. It is necessary to have MOS low-voltage/low-current processes with an EEPROM or FRAM option. These processes provide transistors with low threshold voltages ($V_{TH} \approx 0.5$V) and thus allow complex low-voltage circuitry without need for charge pumps to increase the internal supply voltage.

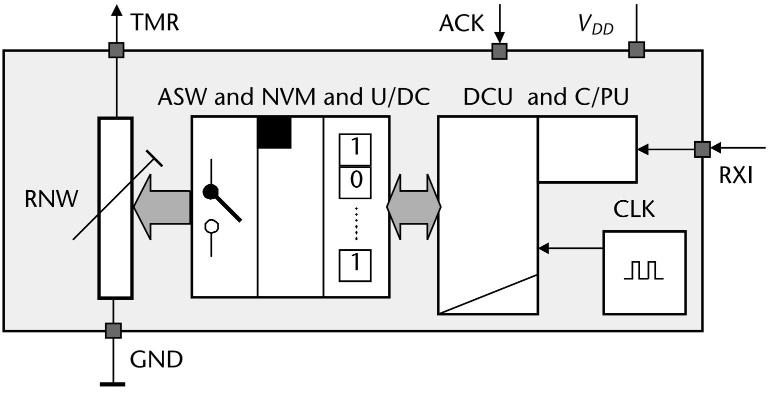


**Figure 12.8** Schematically functional representation of an embodiment of this generic circuit in Figure 12.7 [15].

#### 12.2.2.1 Actuating Unit

Conceptually, an actuating unit (ACU) is a digital-to-analogue resistive convertor, which replaces the adjustment resistor $R_{ADJ}$ (e.g., in Figure 8.34(b)). The sensor features (accuracy, reproducibility, temperature drift, and EMI immunity) are dependent on the quality of this essential part.

The following general features have to be specified: resistive value range (high- or low-impedance), resolution and construction (number of elementary resistors, their value, material, etc.), and the necessary temperature behavior.

The initial question for commitment of a new trimming unit refers to its topology: parallel or series. The answer is already given by the description of the proposed circuit in Figure 12.4. Despite having more benefits, the parallel topology is less flexible and is difficult to be optimized to provide good linearity, and thus, it is not recommended.

In contrast, the series ladder with binary weighted resistors offers good linearity and flexibility in the fixing of the step size. It is a series circuit consisting of more entities (bits), each of these being a parallel connection between a resistive component R and a switching component M (Figure 12.9).

For inductive sensors, a resolution of $m = 6$ to 8 bits should be satisfactory.

The weak point refers to the maximal current allowed to flow through the string. High currents cause high node voltages, particularly when the resistors are not short-circuited by the switches. Especially for the MBS transistor, its source voltage could be so high that the gate voltage is insufficient to neatly switch on it. Hence, the digital ramp with the varying $R_{TMR}$ value could show a striking glitch in the middle of the digital range 0 to 63.

*Example 12.1*

Supposing a series topology (Figure 12.9) with $m = 6$ is decided; the first design step refers to the resistor network (RNW).

Different oscillators presented in Chapter 8 require adjustment resistors $R_{ADJ}$ with a value in the range of the equivalent resistance $R_P$ of the parallel resonant LC-circuit (see (4.38)). Usual values lie in the range $R_P = 1 \ldots 50$ kΩ. To cover this large range with the ASIC and to get a good trimming accuracy, the total resistor $R_{ADJ}$ should be realized with an external preresistor $R_{PRR}$ in series with the ASIC trimming resistor $R_{TMR}$:

$$R_{ADJ} = R_{PRR} + R_{TMR} \tag{12.8}$$

A practical dimensioning formula is

$$R_{TMR\_\max} \geq \frac{R_{P\_\max}}{10} = 5\ \text{k}\Omega \tag{12.9}$$

Due to constructive considerations, a suitable value is $R_{TMR\_\max} = 6.4$ kΩ. Correspondingly, the binary weighted values of the six resistors are:

$$\begin{cases} R_{m-1} = \dfrac{R_{TMR\_\max}}{2} = 3.2\ \text{k}\Omega;\ R_{m-2} = \dfrac{R_{TMR\_\max}}{2^2} = 1.6\ \text{k}\Omega;\ \ldots \\ \ldots R_1 = \dfrac{R_{TMR\_\max}}{2^5} = 0.2\ \text{k}\Omega;\ R_0 = \dfrac{R_{TMR\_\max}}{2^6} = 0.1\ \text{k}\Omega \end{cases} \tag{12.10}$$

and the step size is LSB = 100Ω.

The current limit through resistors string for a typical supply of $V_{DD} = 2.5\text{V}$ is given by:

$$I_{TMR\_\max} \leq \frac{V_{DD}}{2^5 \cdot LSB} = \frac{2500\ \text{mV}}{32 \cdot 0.1\ \text{k}\Omega} = 781\ \mu\text{A} \tag{12.11}$$

and is high enough for oscillators of inductive sensors.

These values suggest a practicable way to integrate the resistors $R_0$ to $R_5$. First, an elementary resistor with a resistance of $R_E = 3.2\ \text{k}\Omega$ is defined and designed. Second, the resistors of RNW result from connecting in parallel corresponding numbers of elementary resistors, which are binary weighted numbers (1, 2, 4, 8, 16, and 32).

Standard integration processes (Section 7.3) offer several types of realizable resistors (at least seven options by bipolar processes and four by CMOS). Technically/economically convenient types are base diffused or ion implanted resistors in bipolar structures and base diffused or polysilicon resistors in CMOS chips [13]. For up-to-date CMOS integrations, the user is well advised to implement polysilicon resistors.

At least one layer of polysilicon (high purity, polycrystalline form of silicon) is required to form the gate of MOS transistors or resistors. The employed geometries are similar to those used for classical diffused resistors and matching properties are also similar. The resistance of an integrated thin rectangular resistor sample is given by [13]:

$$R = \frac{L}{W} \cdot R_{SQ} \tag{12.12}$$

where $L$ and $W$ are the length and width of the rectangular sample. $R_{SQ}$ is the nominal sheet resistance with values in the order of 20 to 80 ohms per square (Ω/sq.). Due to the use of sheet resistance the sample thickness is insignificant.

Considering a typical sheet resistance of $R_{SQ} = 50\ \Omega/\text{sq.}$ and the width of the elementary resistor $W_E = 3\ \mu\text{m}$, its length $L_E$ for $R_E = 3.2\ \text{k}\Omega$ results from (12.12):

$$L_E = \frac{R_E \cdot W_E}{R_{SQ}} = \frac{3200\Omega \cdot 3\ \mu\text{m}}{50\Omega} = 192\ \mu\text{m} \tag{12.13}$$

and the necessary silicon area $A_{SI}$ for the entire RNW (without area for spacings and copper tracks):

$$A_{SI} = W_E \cdot L_E \cdot \left(2^m - 1\right) = 576 \cdot 63\ \mu\text{m}^2 \approx 0.037\ \text{mm}^2 \tag{12.14}$$

is an acceptable value and far smaller than the area on the PCB required by six SMD resistors.

*Example 12.2*

To optimize the temperature behavior of the elementary resistor with $R_E$ = 3.2 kΩ, this device is realized by a series connection of two integrated resistors having different signs of their temperature coefficients. The adequate matching of these components, which is dependent of their temperature coefficients, carries out to a total value of 3.2 kΩ that is theoretically independent of the temperature.

The matching expressions can be determined starting with the general expression of the temperature dependence of an integrated resistor. Its resistance at any temperature $T$ can be expressed by [14]:

$$R(T) = R_a\left[1 + TC1\left(T - T_a\right) + TC2\left(T - T_a\right)^2\right] \tag{12.15}$$

where $R_a$ is the resistance at ambient temperature $T_a$, and $TC1$, and $TC2$ are the linear temperature coefficient [°C$^{-1}$] and the quadratic temperature coefficient [°C$^{-2}$], respectively. $TC2$ is usually negligible, so that the expression becomes linear.

The resistance of the elementary resistor composed of two parts with opposite temperature behaviors becomes:

$$R_E(T) = R_{a+}\left[1 + TC1_+\left(T - T_a\right)\right] + R_{a-}\left[1 + TC1_-\left(T - T_a\right)\right] \tag{12.16}$$

where $R_{a+}$ and $R_{a-}$ are the resistances of the two parts at ambient temperature $T_a$, and $TC1_+$, and $TC1_-$ are the corresponding linear temperature coefficients.

Considering two polysilicon materials with temperature coefficients of $TC1_+$ = +120 ppm/°C and $TC1_-$ = −650 ppm/°C, the factorizing and normalizing of (12.16) gives:

$$R_{a+} = \frac{R_E}{1 + \frac{120}{650}} \text{ and } R_{a-} = \frac{R_E}{1 + \frac{650}{120}} \tag{12.17}$$

For the chosen value of $R_E$ = 3.2 kΩ results: $R_{a+}$ = 2.7 kΩ and $R_{a-}$ = 0.5kΩ.

Generalizing, an RNW with $m$ resistors in series (Figure 12.9) is characterized by the following attributes:

- The network is modular. The elementary resistor has the value of $R_E$ = 3.2 kΩ and almost no temperature dependence.
- The elementary resistor is realized by the series connection of two resistors made from polysilicon with different temperature coefficients and having the following value at room temperature: $R_{a+}$ = 2.7 kΩ and $R_{a-}$ = 0.5kΩ.
- The resistors $R_0$ to $R_{m-1}$ ($m$ = 6 to 8) are obtained connecting in parallel the following numbers of elementary resistors:

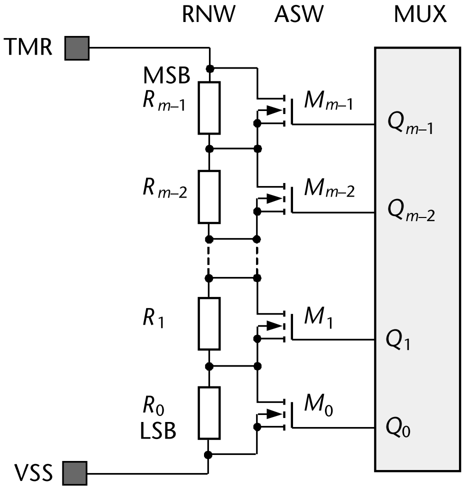


**Figure 12.9** Series topology of an m-bit resistive trimming unit.

$$n_x = 2^{(m-x-1)} \tag{12.18}$$

where $x$ is the resistor level in the network ($x = 0$ to $m - 1$).

The total number of elementary resistors is given by:

$$\sum n_x = \sum_{x=0}^{m-1} n_x = 2^m - 1 \tag{12.19}$$

*Example 12.3*

The switching transistors $M_0$ to $M_5$ in ASW can be realized using integrated n-channel MOS transistors.

The low LSB transistor of 100Ω requires very low channel resistance $R_{ON}$ of the switched on transistors. Usually, this is higher than 100Ω. Hence, a significant reduction of $R_{ON}$ is a condition sine qua non to get a good linearity—and above all—a low temperature drift, especially for the first lower resistance steps.

To meet these requirements, the strategy is to design an elementary switching transistor, which can have a higher $R_{ON}$ but is optimized regarding the temperature drift of $R_{ON}$. The reduction of the $R_{ON}$ value required by the low impedance resistors is achieved connecting in parallel the same binary weighted numbers (1, 2, 4, 8, 16, 32) of switching transistors.

To calculate the $R_{ON}$ and to see how this resistance can be reduced, the computation starts with the complex transmission function of an MOS transistor stated in [14] for the linear region of operation. Under the assumption $V_{DS}$ is low (ideal switch), this expression can be simplified neglecting the second-order term. As a result, the drain current $I_{DS}$ can be expressed by

$$I_{DS} \approx \frac{W}{L} \cdot KP \cdot \left(V_{GS} - V_{TH}\right) \cdot V_{DS} \tag{12.20}$$

where $W$, $L$ are the channel width and length, $KP$ is the transconductance factor (38 $\mu$A/V$^2$, typically), and $V_{TH}$ is the transistor threshold voltage. $V_{GS}$ and $V_{DS}$ are the gate-source and drain-source voltages.

This formula is suitable to give a hand calculation of the $R_{\text{ON}}$:

$$R_{\text{ON}} = \frac{\partial V_{DS}}{\partial I_{DS}} \approx \frac{L}{W} \cdot \frac{1}{KP \cdot \left(V_{GS} - V_{TH}\right)} \tag{12.21}$$

For a typical value of $V_{TH} = 0.6\text{V}$ and due to the low gate-source voltages available by the low-voltage ASIC supply with $V_{DD} = 2.5\text{V}$ (two-wire application), $R_{\text{ON}}$ has high values, even for an impressive $L/W$ ratio (e.g., 1:100):

$$R_{\text{ON}} \approx \frac{1}{100} \cdot \frac{1}{38 \cdot 10^{-6} \cdot 1.9} = 138.5\Omega \tag{12.22}$$

The conclusion is unambiguous; the switches of the ASW must be realized connecting in parallel such elementary MOS transistors. Their number has to be equal to those of the resistors (see above) or higher. The experience shows the fact that the hand calculation only gives an estimation.

The PSpice simulations of the entire ACU unit are more than required in this case. Only by means of this tool is it possible to make a real computation of the network parameters (linearity, accuracy, step-size constancy, monotonicity, temperature drift, etc.).

#### 12.2.2.2 Control Unit and Peripherals

The control unit CU/P manages the entire ASIC operation, which essentially consists of the decoding of the programming commands sent by the programming unit and their corresponding execution. Optimally, it is a hardware-wired Mealy state machine, which manages a programming flow as typically represented in Figure 12.10.

The clock generator CLK is essentially an astable multivibrator with an integrated capacitor that is charged and discharged by a current source and a current sink. Much easier, the multivibrator could be an RC oscillator. The stage contains simple 3-bit binary frequency dividers and provides three clock signals:

1. *CLOCK0* as a main clock for the DCU and with the frequency $f_0$ (e.g., ≤64 kHz);
2. *CLOCK1*, which is a clock signal for the C/PU and has the frequency $f_1 = f_0/8$;
3. *CLOCK2,* acting as a clock for the U/DC and having the frequency $f_2 = f_0/64$.

The power-on reset (POR) stage is a classical monostable multivibrator and generates a single reset pulse if the supply voltage $V_{SS}$ is applied and reaches a fixed threshold voltage (e.g., the bandgap reference). The pulse complies with the standards and has a duration of 10 ms, typically.

#### 12.2.2.3 Tech-In Trimming

The tech-in trimming strategy is based on the unambiguous relationship between the amplitude of the oscillator output signal $V_{lcm}$ and the value of the adjustment resistor $\mathrm{R}_{\mathrm{ADJ}}$, which characterizes all oscillator circuits for inductive sensors described in Chapter 8. Hence, an increase of $\mathrm{R}_{\mathrm{ADJ}}$ starting with a minimal value provokes a continuous decrease $V_{lcm}$ or vice versa, theoretically.

During the trimming, the signal $V_{lcm}$ is compared in the sensor electronics with a threshold voltage by the comparator CMP (Figure 7.1). If $\mathrm{R}_{\mathrm{ADJ}}$ achieves a value, for which $V_{lcm}$ equalizes the threshold, the comparator changes the state and its output can be used as an acknowledge signal ACK for the trimming ASIC to stop the trimming.

Hence, the trimming operation needs only to be launched by the operator. It runs fully automatically and stops if the adjustment is performed [15]. This aspect distinguishes the method from traditional versions in which the trimming is step-by-step conducted by the user or needs to be stopped by the operator.

To maintain compliance with the standards, the sensor with trimming ASIC has no additional connection terminals. The programming commands will be transmitted from the programming unit to the sensor modulated on the DC supply voltage. In the sensor, they will be extracted from the supply line and fed to the ASIC input RXI. This channel is latently existent (but not used) during normal operation, when the sensor is supplied from industrial networks, which could be strongly disturbed. Hence, the path should have a high interference immunity. Protection measures are described in Section 12.2.2.4.

The trimming starts (Figure 12.10) with the mechanical adjustment, provided by the user, of the wanted distance between the standard target or specific metallic object and the active face of the inductive sensor with the ASIC (see Figure 2.2).

The supply lines of the sensor with programming features are temporarily connected to the output terminals of the programming unit, which generates power supply and programming commands. If this power voltage is switched on, the initialization phase occurs in the ASIC and the device is prepared for trimming. This waiting state should be time-limited by a time window to increase the system safety.

If the user manually releases a programming instruction by touching a button on the programming unit, this command is executed in the ASIC. The incrementing or decrementing of $\mathrm{R}_{\mathrm{ASW}}$ is preceded by a reset or set of $\mathrm{R}_{\mathrm{ASW}}$ and automatically runs until the stop signal ACK arrives from the sensor electronics, indicating the achievement of the wanted distance. The speed is fixed, imposed by the CLK. Instant commands (reset or set of $\mathrm{R}_{\mathrm{ASW}}$) without running are available for test tasks.

At the end, the digital trimming value is stored in the NVM and the ASW is reconnected to NVM (normal operation path).

Using multitime programmable memories (EEPROM or FRAM), the request for repeatable programming cycles is fulfilled for a high number of tries.

#### 12.2.2.4 Communication and Programming Unit C/PU

The ASIC interface to receive programming commands is based on a concrete communication method [15] successfully used on a large scale for 20 years [16, 17]. It

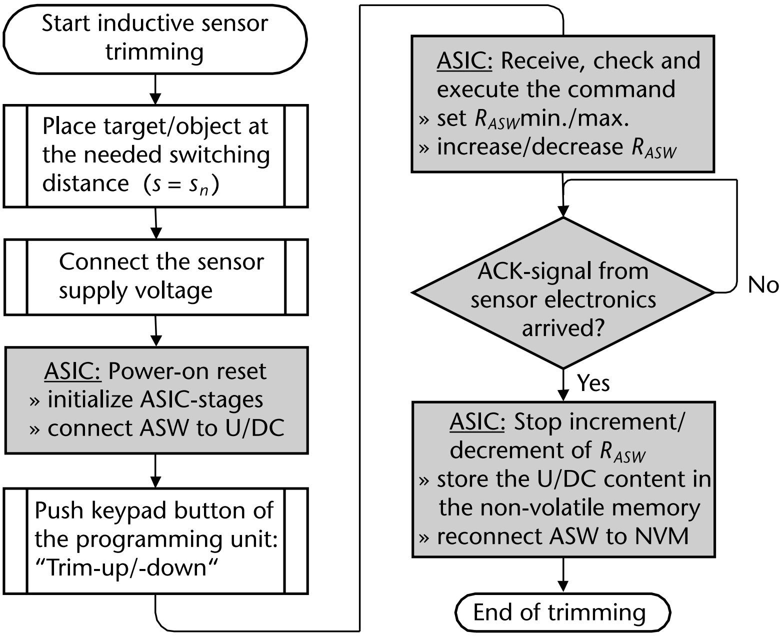

Start inductive sensor trimming
Place target/object at the needed switching distance ($s = s_n$)
Connect the sensor supply voltage
ASIC: Power-on reset
» initialize ASIC-stages
» connect ASW to U/DC
Push keypad button of the programming unit: "Trim-up/-down"
ASIC: Receive, check and execute the command
» set $R_{ASW}$min./max.
» increase/decrease $R_{ASW}$
ACK-signal from sensor electronics arrived?
No
Yes
ASIC: Stop increment/ decrement of $R_{ASW}$
» store the U/DC content in the non-volatile memory
» reconnect ASW to NVM
End of trimming

**Figure 12.10** Flowchart of a trimming procedure with the trimming ASIC (gray shaded blocks represent the ASIC tasks and results).

provides high data protection and can be used in industrial environments with high electromagnetic disturbances.

The communication between the programming unit and the trimming ASIC is asynchronous and unidirectional and occurs following a proprietary protocol. However, the communication is in accordance with the ISO/OSI model and implements the first two layers (see Section 14.1.4):

a. The Data Link Layer, which realizes the safe data transmission. The communication protocol is similar to the UART described in Section 14.1.3.2 and uses one or two 8-bit frames, with the following bit-fields:
   - *START*-bit (INST<0>).
   - *OPEN KEY*-bits (INST<1:2>) that facilitate a reliable recognition of the telegram and activate the path to C/PU.
   - *DATA* FIELD-bits (INST<3:5>), containing instruction codes for the trimming ASIC. The four instructions presented in Figure 12.10 represent a minimal set of instructions. The set can contain additional various instructions and when the 3 bits are insufficient, a second telegram can be implemented. Nonassigned bit combinations are ignored.
   - *PARITY CHECK*-bit (INST<6>) increases the communication security, providing the lowest Hamming distance $d = 2$ (Section 14.1.3.1).
   - *STOP*-bit (INST<7>).

b. The Physical Layer realizes the data transmission on the bit level and can be characterized by the following features:
   - The physical medium for the transmission is the two sensor supply voltage terminals wired to the output of the programming unit;

- The carrier signal is a sinusoidal wave superposed on the supply voltage;
- The carrier modulation is a 100% AM (amplitude modulation);
- The bit coding is MANCHESTER II (see Section 14.1.3.3) with times imposed by the signal CLOCK1 (Section 12.2.2.2).

To summarize, procedures that are similar to the above-described methods and have parameters (voltages, frequencies, times) adapted to the sensor attributes (voltage limitations, low-pass filter behaviors, etc.) offer a high level of data transmission safety and immunity to disturbances with a reasonable need of hardware, meaning chip area consumption and/or and external components as quartz crystal XTAL, and so forth. Errored command decodings, which could parasitically launch a false programming, memory data loss, or faulty data storage are the most dangerous events and should be strictly avoided.

### 12.2.2.5 Improved ACU units

The local temperature compensation of the high-impedance elementary resistor ($R_E = 3.2$ kΩ) (Section 12.2.2.1) makes sense if the trimming ASIC is an autarkic device and its chip temperature cannot be matched with the temperature of the inductive sensor electronics.

When the trimming network is monolithically integrated with the sensor electronics, the elementary resistors could be uncompensated and the temperature compensation occurs in the monolithic device. Instead of a high-impedance elementary resistor and a top-down resistance reduction, the strategy uses a low-impedance elementary resistor and a bottom-up resistance increase occurs by series connections.

*Example 12.4*

Maintaining the series topology for the RNW and the same resolution ($m = 6$ and LSB = 100Ω), the resistor series string consists now of 63 elementary resistor of 0.1 kΩ (Figure 12.11(b)).

In addition to this modification, Figure 12.11(a) shows two schematic extensions, which carry to the following technical perfections:

- Improvement of the trimming resolution by parallel connecting several identical resistor series strings (Figure 12.11(b)), which are synchronously varied (common mode). For a total number of $L = 8$ strings, the resulting step size drops to 100Ω: 8 = 12.5Ω. Unfortunately, this enhancement unfavorably reduces the trimming range of $R_{TMR}$ to 6.3 kΩ: 8 = 0.788Ω;
- Enlargement of the total trimming resistance by the integration of an additional series network of preresistors $\mathrm{R}_{\mathrm{PRR}}$ (see (12.8)). For a total number of $K = 8$ segments of each $X = 3.2$ kΩ, the resulting preresistor can be programmed in the range:

$$R_{PRR} \in [0, 25.6 \text{ k}\Omega] \tag{12.23}$$

Example 12.4, with its setup: $K = 8$, $L = 8$, $X = 3.2$ kΩ, and $R_{\mathrm{ON}} = 135\Omega$ represents a particularly good practical example.

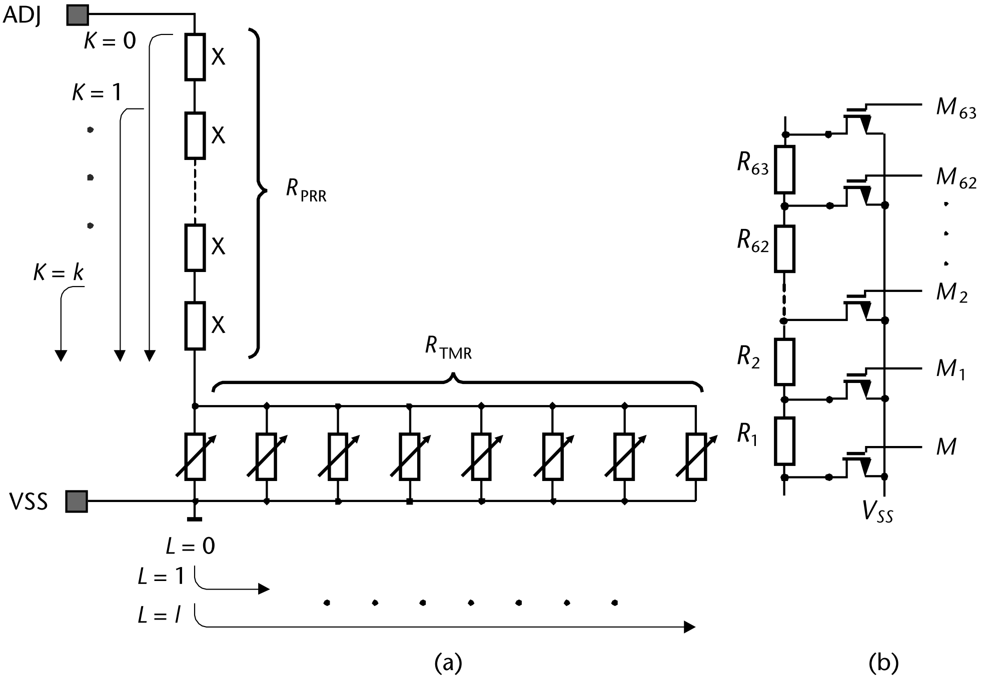


**Figure 12.11** (a) Complex trimming resistor network allowing the improvement of the resolution and/or inserting of a programmable preresistor $R_{PRR}$, and (b) detail of an $R_{TMR}$-segment.

In conclusion, the complex network in Figure 12.11 is a flexible and universally applicable trimming network for all possible LC oscillator variants implemented in inductive sensors, including these with many critical PCB coils (Section 5.4) [18].

At any time, the setup can be adapted to the concrete sensor demands. The total resistance can be calculated using the general formula:

$$R_{ADJ} = R_{PRR} + R_{TMR} = X \cdot K + (T + 1) \cdot \frac{0.1\ \mathrm{k\Omega}}{L + 1} \tag{12.24}$$

where

$K$ is the variable, which defines the magnitude of the preresistor. This is to be set by a direct command having an UART frame similar to those for standard commands (Section 12.2.2.4).

$X$ is the specific step size of the preresistor.

$L$ is the variable that determines the granularity of the trimming and has to be directly programmed by another similar command.

$T$ is the current variable, which automatically runs during the teach-in and is stopped to the sensor specific trimming value if the acknowledgment is available (Figure 12.10).

## References

[1] Bluhm, T., *Using Laser Trimmable Resistors*, Vishay Beyschlag Application Note, https://www.vishay.com/docs/28893/usinglasertrimmableresistors.pdf, 2013.

[2] https://datasheet.octopart.com/111N-Microbridge-Technologies-datasheet-8331023.pdf, MBD-153-AS Dual Adjustable 15KΩ–10KΩ Low-TCR Rejustor™, Microbridge Technologies, Inc., 2008.

[3] http://data.leocom.kr/datasheets/202763_MBK-500-x00_Matchbox_DS_V1_1.pdf, Rejustor™ Matchbox MBK-500-x00, Microbridge Technologies, Inc., 2007.

[4] https://www.bourns.com/docs/technical-documents/technical-library/sensors-controls/publications/SC_Solutions_Guide.pdf, *Sensors & Controls Solutions Guide*, Bourns, Inc., 2017.

[5] http://www.analog.com/media/en/news-marketing-collateral/product-selection-guide/precision-technology-selection-guide.pdf, catalog, *Precision Technology Selection Guide*, Analog Devices, 2018.

[6] http://www.analog.com/media/en/technical-documentation/data-sheets/AD5291_5292.pdf, AD5291 256-Position, Digital Potentiometers with Maximum ±1% R-Tolerance Error and 20-TP Memory, Analog Devices, 2012.

[7] http://www.analog.com/media/en/technical-documentation/data-sheets/AD5259.pdf, data sheet, AD5259 Nonvolatile, $I^2C$-Compatible 256-Position, Digital Potentiometer, Analog Devices, 2012.

[8] Fericean, S., F. Kruepl, M. Fritton, and T. Reider, Electronic Component for a Sensor Apparatus, Sensor Apparatus and Method of Configuring a Sensor Apparatus, U.S. Patent 8,660,806, issued February 25, 2014.

[9] Klemm, T., and M. Hamma, Digital geht's besser, Elektronik Fachzeitschrift, www.elektroniknet.de, August 22, 1995.

[10] Hering, E., G. Schönfelder, *Sensoren in Wissenschaft und Technik—Funktionsweise und Einsatzgebiete*, 2nd Edition, Wiesbaden, Germany: Vieweg+Teubner Verlag/Springer Fachmedien GmbH, 2018.

[11] Westrup, K.- P., Inductive Proximity Switches with Electronic Adjustment, DE 102013218405B4, March 30, 2017.

[12] Gundlach, J., et al., Electronically Adjustable Inductive Proximity Switches, DE 102015212412A1, January 5, 2017.

[13] Gray, P. R., P. J. Hurst, S. H. Lewis, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, Fifth Edition, New York: John Wiley & Sons, 2009.

[14] Vladimirescu, A., *The SPICE Book*, New York: John Wiley & Sons, 1994.

[15] Fericean, S., H. Kammerer, and H.- W. Plank, Proximity Switch Operating in a Non-Contacting Manner, U.S. Patent 5,408,132, issued April 18, 1995.

[16] Fericean, S., M. Friedrich, M. Fritton, and T. Reider, *Moderne Wirbelstrom-Sensoren–linear und temperaturstabil*, Elektronik Fachzeitschrift, www.elektroniknet.de, April 2001.

[17] Jagiella, M., S. Fericean, "Miniaturized Inductive Sensors for Industrial Applications," *Proceedings/First IEEE Conference on Sensors*, Orlando, FL, 2002.

[18] http://www.ti.com/lit/an/snoa954b/snoa954b.pdf. Chris Oberhauser, SNOA954 Application Report, LDC Device Selection Guide, Texas Instruments, July 2016.

CHAPTER 13

# Inductive Sensors: Temperature Compensation

This chapter presents various methods for the compensation of the temperature dependence of an inductive sensor.

An efficient compensation requires the permanent measuring of the sensor temperature. For this reason, the chapter starts with a first section, which deals with conventional temperature probes and measuring procedures that provide more accuracy levels of the temperature gauging.

## 13.1 Temperature Sensing Devices

Temperature sensing devices constitute an important part of inductive sensors and are used as a discrete device or integrated in a sensor ASIC. The integrated temperature sensor is based on the bipolar transistor. Diodes and integrated resistors are also sometimes used for convenience. Transistors are adequate within a temperature range of −50°C to +180°C. In silicon technology, the absolute resistance values of integrated resistors (e.g., in Section 12.2.2.1) are usually not very accurate (±20%, typically), but the matching of the ratio between two resistances is very good (±0.1%).

Thin-film resistors can be used as a second option in wider operating temperature ranges than integrated transistors, due to the absence of the p-n junctions that give bad electrical isolation at high temperatures. They are made of polysilicon or platinum.

As discrete elements, thermistors or platinum resistors are used quite often.

### 13.1.1 Passive Temperature Probes

Passive temperature probes are elementary sensors where the temperature is measured not by means of volume or length changes but from the temperature dependence of the electrical resistance. They are also called metal-resistor thermometers.

#### 13.1.1.1 Negative-Temperture and Positive-Temperature Coefficient Thermistors

Negative-temperature coefficient thermistors (NTCs) and positive-temperature coefficient thermistors (PTCs) [1] are often used in inductive sensors, particularly in those belonging to the levels I and II (see Figure 7.2). The thermistors have an ordinary structure, their manufacturing costs are very low, and the implementation is immediate.

NTCs are applied for temperature compensation of sensor features. PTCs are preponderantly used for overload current protections of the sensors outputs. A typical application is shown in Figure 7.15. The collector of $T_3$ forms the effective switching output of the sensor (see Section 10.2). A POSISTOR PTC thermistor (manufacturer Murata Manufacturing Co.) [2] is inserted between this collector and the real sensor output OUT and offers a simple but efficient overcurrent protection. Under normal operation, its resistance has a negligible value (below 1Ω). If the output current exceeds the specified values, the PCT temperature rises up above 150°C and the PTC resistance increases to above 10 kΩ, which limits the current.

NTCs are semiconductor ceramic or polymer resistors with a significant negative temperature coefficient (≈ −2 to −5%/°K) and a large temperature range of −55 to +330°C. The complex conducting mechanism carries to an exponential temperature dependence of the NTC resistance (Figure 13.1). The equation of the curve in Figure 13.1(a) can be expressed by:

$$R(T) = R_N \cdot e^{B\left(\frac{1}{T}-\frac{1}{T_N}\right)} \tag{13.1}$$

where $R(T)$ is the NTC resistance at the temperature $T$, $R_N$ is the nominal resistance at the nominal temperature $T_N$ (temperatures in kelvin), and the constant $B$ is the $B$-value of the NTC, defined as:

$$B = \frac{E_a}{k} \tag{13.2}$$

where $E_a$ is the activation energy and $k$ = Boltzmann's constant (see (8.103)). The achieved accuracy with this formula is ±1°C over the range of 0°C to +100°C.

The best approximation known today is the Steinhart-Hart formula, published in 1968 [3]:

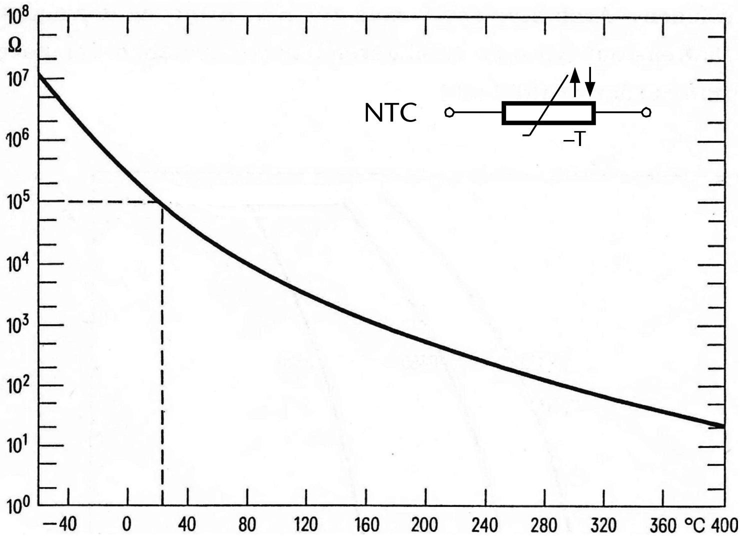
$$\frac{1}{T} = \frac{1}{T_N} + \frac{1}{B}\ln\frac{R(T)}{R_N} + \frac{1}{C}\left[\ln\frac{R(T)}{R_N}\right]^3 \tag{13.3}$$

The formula (13.3) contains a new term with the coefficient $1/C$.

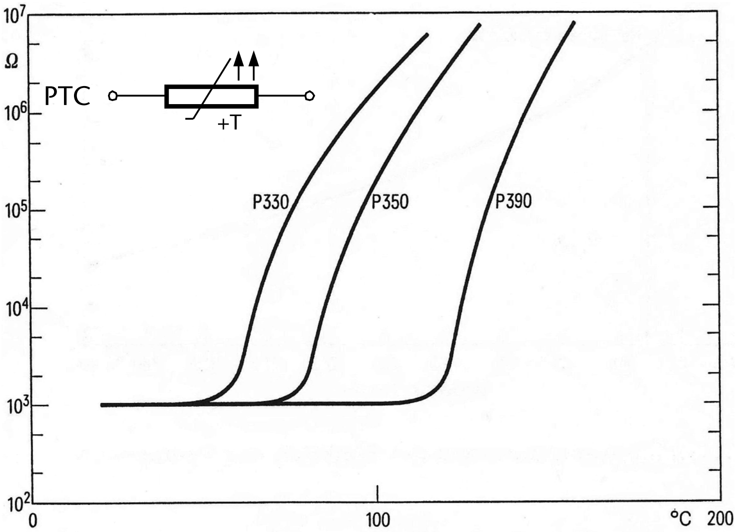
The industry offers NTCs with $R_N$ = 1Ω to 10 MΩ and tolerances from ±20% to ±5% or ±2%. The $B$ value is a specific material constant and lies in the range: 1500°K to 7000°K.

In a large spectrum of applications, the NTC is used as an element for the temperature compensation of oscillators or for the stabilization of the transistor operating point.

In general, the temperature coefficient of the output signal $V_{lcm}$ of an oscillator for inductive sensors has a negative value. The easiest method to keep this voltage constant is to increase the feedback current in the oscillator with rising temperature by decreasing the value of the adjustment resistor $R_{ADJ}$. This can be done by replacing the resistor $R_{ADJ}$ with a parallel-series network consisting of an NTC in parallel with a resistor, still connected in series with a second matched resistor.



(a)



(b)

**Figure 13.1** NTC and PTC resistance dependencies on the temperature and their circuit symbols.

This method is applicable for all oscillator circuits described in Chapter 8 (e.g., Figure 8.21(b)). Thus, the NTC has a particular importance for inductive sensors.

If the NTC is used not as compensation device but as a temperature probe, its $R(T)$ characteristic can be linearized connecting a temperature-independent resistor $R_{LIN}$ in parallel to NTC. The resistance $R_{NW}$ of the linearized network results in:

$$R_{NW} = \frac{R(T) \cdot R_{LIN}}{R(T) + R_{LIN}} \tag{13.4}$$

The choice of the $R_{LIN}$ is made in such a manner that the curve $R_{NW}(T)$ gets an inflexion point in the middle of the measurement range.

The PTCs are resistors made by sintering methods from barium-titanate (BaTiO3). They have a high positive temperature coefficient (≈ +7%/°K to +70%/°K) in a temperature range that can be technologically defined. The resistance at moderate temperatures $R(T)$ can be expressed by

$$R(T) = R_N \cdot e^{A(T-T_N)} \tag{13.5}$$

where $R_N$ is the nominal resistance at the nominal temperature $T_N$, while $A$ is the temperature coefficient (e.g., 0.16/°K) [4]. Note that this value is 40 times higher than that of metal resistors. The tolerance range of the $R_N$ is ±10%.

The PTC voltage-current characteristic has an incipient linear rising segment. If the voltage increases, PTC temperature slowly rises until the activating temperature is reached. In this moment, the resistance grows, the current goes down, and an equilibrium state is installed. The internal heating effect carries out to growths of the resistance by a factor of 1,000 or high as a result of a high power dissipation. The activating temperature (40°C to 180°C) is specified within a narrow range (±10%).

This PTC protection property was intensively used in inductive sensors belonging to older generations. It is a very safe and comfortable solution from the point of view of sensor manufacturers. In contrast, the solution was accepted but not agreed on by sensor customers due to the considerable and fluctuating voltage drop between the supply voltage and the sensors output and because of the large necessary recovery time after a short circuit.

#### 13.1.1.2 Nickel and Platinum Metal-Resistor Thermometers

The discrete nickel and platinum resistors are well known for their range, accuracy, and stability, serving as a temperature standard over a wide temperature range.

They are basically a thin-layer chip resistor consisting of thin nickel or platinum meander-shaped layer, which is printed on a ceramic substrate (by photolithography or laser forming) and is followed by a protecting passivation. Both versions provide 100, 500, and 1,000Ω at 0°C. The resistance dependencies on the temperature are unfortunately nonlinear.

Nickel resistor thermometers are less expensive and have higher temperature coefficients. The standard DIN 43 760 states the temperature range of −60°C to +250°C and defines their temperature behavior by means of a polynomial of grade $n \geq 6$:

$$R(T) = R_0 \cdot \left(1 + a \cdot T + b \cdot T^2 + c \cdot T^4 + \ldots\right) \tag{13.6}$$

where $R_0$ is the nominal resistance at the nominal temperature 0°C and the linear temperature coefficient is $a = 6.18 \cdot 10^{-3}$/°K.

The platinum resistor thermometers are stated in the standard DIN IEC 751, which defines two different temperature ranges: −200°C to 0°C and 0°C to +850°C and two corresponding polynomials. For the range of positive temperatures, the expression (13.6) is valid and $a = 3.805 \cdot 10^{-3}$/°K.

The classical platinum thermometer is the device Pt100. It has a resistance value of 100Ω at 0°C and is available in many qualities, the basic quality being compliant to the IEC 751 standard, with a temperature dependence of about 0.38%/°K and a tolerance of ±0.15°K at 0°C. Better qualities (class A) are available at higher prices. The linearity of Pt100 in the range of 0°C to +250°C is very good. Thus, Pt100 is an ideal sensor for the temperature ranges of industrial sensors.

Other resistance values are also used, for instance Pt1000, for low measurement currents. The greater the nominal value, the higher the sensitivity to temperature changes.

The metal-resistor thermometers are passive elementary sensors and require an external power supply. In addition, to achieve high accuracy, the four-wire connection technique (two I-wires for the supply current of Pt100 and two V-wires to perform a high impedance voltage measuring) has to be used. There is no current through the measuring lines, and thus, no parasitic voltage drops.

*Example 13.1*

Based on our knowledge gained in the previous chapters, a performant supply and evaluation unit for Pt100 can now be designed. The unit should consist of

- An adjustable voltage-controlled current source with op amp (see Section 9.1) to supply along the first wire-pair;
- A high performance instrumentation amplifier (version with three op-amps in Figure 9.3) to catch the voltage sourced by Pt100 along the second wire-pair;
- An additional linearization stage, if needed (see Section 9.4);
- An additional circuit, which uses a 4- to 20-mA current output stage (see Section 10.1.2) to make current telemetry possible.

The architecture proposed above is only an embodiment. The reader is invited to create more versions.

As already anticipated in Section 10.1.2, a similar unit is based on the device AD693 [5], which is described in [6]. This is a loop-powered, 4- to 20-mA sensor transmitter with three stages: (1) an instrumentation amplifier front-end, (2) the V/I converter, and (3) a reference voltage. It performs a monolithic signal conditioning of low-level inputs from a variety of transducers to control a standard 4- to 20-mA, two-wire current loop and has a large range of applications. One of these refers to the device Pt100, described as a 100Ω platinum resistance temperature detector RTD (see [5], pp. 9, 10). The IC is fab-out precalibrated for the correlation of 4- to 20-mA full output span to 0°C to 104°C (exactly for $a = 3.805 \cdot 10^{-3}$/°K). In addition, five other precalibrated ranges are selectable by two jumps between specific pins. The IC has several hardware options to provide other settings.

### 13.1.2 Active Temperature-Dependent Circuits

#### 13.1.2.1 Bipolar Transistor Temperature Sensors

The temperature sensing capability of a transistor depends on the behavior of the silicon p-n junction [7]. To use the transistor as a temperature sensor, it is reasonable to operate this in two-terminal configuration (Figure 13.2.a) with short-circuited base-collector terminals to prevent base-width modulation (Early effect) due to the changes in the collector-base voltage.

When this circuit is supplied by constant current $I_C$, the base-emitter voltage $V_{BE}$ decreases almost linearly with the temperature:

$$V_{\mathrm{BE}}(T) = V_{\Phi} - \alpha \cdot T \tag{13.7}$$

where $\alpha$ is a sensitivity coefficient (2 mV/°K, typically), which depends on the bias-current density and process parameters, and $T$ is the absolute temperature (expressed in kelvin). The constant $V_{\Phi} \approx 1.25\text{V}$ is the bandgap voltage and denotes the base-emitter value at 0°K (–273°C). This value is process specific but independent of the bias-current or transistor geometry [8]. Thus, the sensor characteristic is nearly a straight line with a negative slope ($\approx -2\text{mV/°K}$).

As a result, the circuit in Figure 13.2(a) is quite suitable for temperature measuring. With a single-point calibration (i.e., the measurement of $V_{BE}$ at an arbitrary temperature, the complete straight line $V_{BE}$ $(T)$ can be known over a wide temperature range). Usually, the voltage $V_{\Phi}$ is precisely known for the used fabrication process and the calibration point is chosen at the temperature $T^*$ at which $V_{BE}$ $(T^*)$ $= V_{\Phi}/2$. This temperature is closed to the room temperature.

To determine the current dependence of $V_{BE}$, the already used expression of the collector current (8.110) can be changed to:

$$I_{\mathrm{C}} = A_E \cdot I_S \cdot e^{\frac{V_{BE}}{V_T}} \tag{13.8}$$

where $A_E$ is the emitter area, $V_{BE}$ the base-emitter voltage, $I_S$ the saturation current, and $V_T$ is the thermal voltage:

$$V_T = \frac{k \cdot T}{q} \approx 26 \text{ mV at } 300\text{°K} \tag{13.9}$$

with

$k$ = Boltzmann's constant ($1.380 \cdot 10^{-23}$ joules/°K)

$q$ = the electron charge ($1.602 \cdot 10^{-19}$ coulombs);

$k/q$ = 86.17 μV/K

$T$ = the absolute temperature in kelvin

By taking the natural logarithms of the current $I_C$ in (13.8), the $V_{BE}$ results in:

$$V_{BE} = \frac{k \cdot T}{q} \cdot \ln \frac{I_C}{A_E \cdot I_S} \tag{13.10}$$

As mentioned above, it is sufficient to trim the transistor at a single temperature to make the transistor characteristic equal to the nominal one. Because of production tolerances, there is a sample deviation from the nominal temperature characteristic. The sensors can be calibrated by adjusting the biasing current or the emitter area.

According to (13.10), changing $I_C/A_E$ by a factor of 2 modifies the $V_{BE}$ by $V_T \cdot \ln 2 = 26\ \text{mV} \cdot 0.693 = 18\ \text{mV}$.

Integrated transistors, used as temperature probe, could have a spread in $V_{BE}$ of about 6 mV (standard deviation) [7]. The required range of the collector current needed to adjust this deviation is ±6 mV/26 mV ≈ ±23%.

The behavior can be significantly improved applying a current feedback by means of a current mirror (Figure 13.2(b)). The current $I_{C2}$ is not set anymore by a resistor but is derived from a reference current $I_R$. If the transistors are matched and the input-output transfer ratio of the current mirror $T_3 - T_4$ is 1:1, the currents are equal.

Considering (13.10), the reference current meets the transcendental equation:

$$I_R \approx \frac{V_T}{R_1} \ln\left(\frac{I_R}{I_{S2}}\right) \tag{13.11}$$

The solution of this equation depends on $V_T$, $R_1$ and $I_{S2}$ but not on the supply voltage (as long as the Early effect can be neglected). The current $I_{C2}$ is also stabilized, thus the base-emitter voltage of $T_2$ could be considered constant and the following equation is valid:

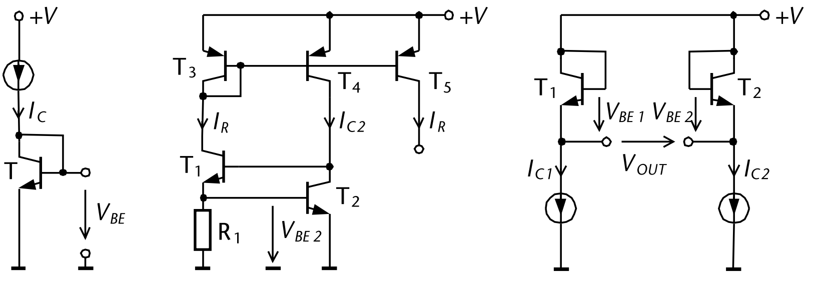
$$I_R \approx \frac{V_{BE2}}{R_1} \tag{13.12}$$

To extract the current with a negative temperature coefficient, the second output transistor $T_5$ of the double current mirror is used as a current source with the same current.

The accuracy can be improved to the detriment of the sensitivity using a dual circuit, which consists of two matched transistors exposed to the same temperature and supplied by two DC current sources (Figure 13.2(c)). The collector currents or/and the emitter areas could be different.

If collector currents are different and the emitter areas are equal, based on (13.10), the output voltage results in:

$$V_{\text{OUT}} = V_{BE1} - V_{BE2} = \frac{k \cdot T}{q} \cdot \ln\frac{I_{C1}}{I_{C2}} = 86.17\ \frac{\mu\text{V}}{^\circ\text{K}} \cdot \text{T} \cdot \ln n \tag{13.13}$$



(a) (b) (c)

**Figure 13.2** Single-path and dual-path VBE temperature sensors.

where $n$ is the ratio of the collector currents.

If collector currents are equal but the emitter areas are unequal, the expression of the output voltage is identical to those in (13.13); the only difference is the reversed expression of $n = A_{E2}/A_{E1}$.

The derivative of $V_{\text{OUT}}$ with respect to $T$ offers the circuit sensitivity:

$$S \overset{\text{def}}{=} \frac{dV_{\text{OUT}}}{dT} = 86.17\ \frac{\mu\text{V}}{°\text{K}} \cdot \ln n \qquad (13.14)$$

which is noticeably lower than those of the single transistor version. Even for a high value $n = 10$, the sensitivity is $S = 0{,}199$ mV/°K (tenth part of 2 mV/°K).

### 13.1.2.2 PTAT Sensors

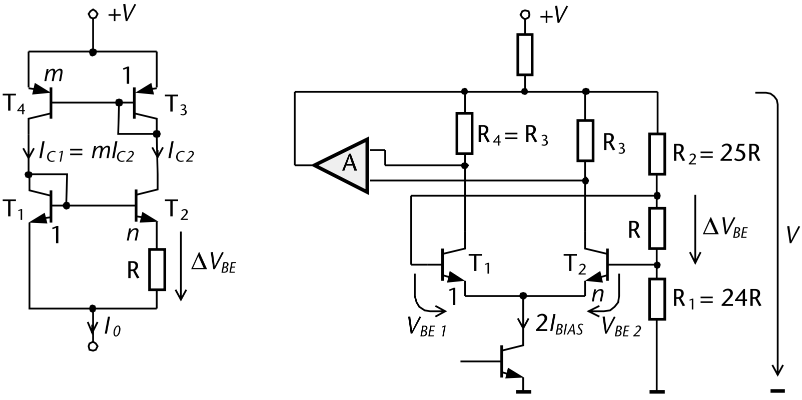
The proportional to the absolute temperature (PTAT) sensors are the best-known temperature sensing circuits, and can be integrated on the same chip together with the specific stages of inductive sensors.

The PTAT current source shown in Figure 13.3(a) generates a current proportional to the absolute temperature. The basic signal is the difference $\Delta V_{BE}$ between the base-emitter voltages of two transistors $T_1$ and $T_2$ operated at a constant ratio of their emitter-current densities. When the transistors are at the same temperature and their current amplification factor $\beta$ is high enough, to consider the emitter current equal to the corresponding collector current, (13.10) gives:

$$\Delta V_{BE} = V_{BE1} - V_{BE2} = \frac{k \cdot T}{q} \cdot \ln \frac{I_{C1} \cdot A_{E2} \cdot I_{S2}}{I_{C2} \cdot A_{E1} \cdot I_{S1}} \qquad (13.15)$$

For identical transistors fabricated on the same chip, it holds that $I_{S1} = I_{S2}$.

In the general case, the transistors have different emitter areas and the ratio is $n = A_{E2}/A_{E1}$. When the current mirror $T_3 - T_4$ keeps a constant collector-current ratio $m = I_{C1}/I_{C2} \neq 1$, the final expression is:



(a) (b)

**Figure 13.3** PTAT temperature sensors: (a) current source, and (b) voltage source.

$$\Delta V_{BE} = \frac{k \cdot T}{q} \cdot \ln(n \cdot m) \tag{13.16}$$

This voltage is a PTAT voltage. Usually, the multiplication factor is chosen equal to 8 (ln8 ≈ 2) and the slope of the linear characteristic is:

$$S \overset{\text{def}}{=} \frac{\Delta V_{BE}}{dT} = 86.17 \ \frac{\mu\text{V}}{°\text{K}} \cdot \ln 8 \approx 0.179 \ \frac{\text{mV}}{°\text{K}} \tag{13.17}$$

The PTAT current source provides an output current which is PTAT and stabilized against changes in the supply voltage. If $m = 1$ and $n = 8$, the collector currents are equal $I_{C1} = I_{C2}$ and the PTAT current can be expressed by:

$$I_0 = \frac{2}{R} \cdot \frac{k \cdot T}{q} \cdot \ln 8 \tag{13.18}$$

A typical resistor value is R = 360Ω. It adjusts a current source sensitivity of 1 μA/°K.

A drawback of the PTAT sources is the presence of a large offset signal at ordinary temperatures. To have a temperature signal with its zero in the range of interest, a differential amplifier (Section 9.1) has to be used. It calculates the difference between the initial PTAT signal and a reference signal and delivers the final signal.

The second version in Figure 13.3(b) consists of a differential amplifier, which has a base loop consisting of two BE junctions with very different emitter areas (1 to 10) and a series resistor R. The collector currents of $T_1$ and $T_2$ have to be equal. The subsequent amplifier A controls this equality readjusting the main voltage to fulfill this requirement. The voltage across the resistor R can be expressed by:

$$V_R = V_{BE1} - V_{BE2} = \frac{k \cdot T}{q} \cdot \ln 10 \approx 0.2 \ \frac{\text{mV}}{°\text{K}} \cdot T \tag{13.19}$$

where $T$ is the absolute temperature in kelvin.

The voltage $V_R$ drives the current through the resistor R and, hence, through the full voltage divider $R_2$, R, $R_1$.The total voltage drop becomes:

$$V_T = \left(R_1 + R + R_2\right) \cdot \frac{V_R}{R} = 50R \cdot \frac{V_R}{R} = 50 \cdot V_R \tag{13.20}$$

Substituting (13.19) in (13.20), an elegant expression of the PTAT voltage results:

$$V_T^{[\text{mV}]} = 10 \cdot T^{[°\text{K}]} \tag{13.21}$$

This basic PTAT voltage is usually amplified and buffered.

Both schematics are implemented; the first version is preferably used in specific ASICs or in some temperature transducers (e.g., AD590), whereas the version with feedback amplifier is met in commercial temperature sensor ICs. There are more market leaders that offer integrated temperature sensors.

A remarkably wide variety of integrated precision Fahrenheit, Celsius, and Kelvin temperature sensors (more than 30 types, analog or digital) is singly offered by Texas Instruments [9]. Table 13.1 summarizes some of these devices.

### 13.1.3 Active Temperature-Independent References: Bandgap References

Aside from devices or circuits, which very accurately measure the ambient or sensor internal temperature, providing a signal with an exactly defined temperature dependence (actual-value indicators), the sensor temperature compensation requires reference units whose outputs are constant and theoretically temperature-independent (nominal-value indicators).

These high-precision references have a common attribute. They are called bandgap references because they generate a constant, temperature-independent voltage whose value is more closed to the bandgap voltage ($V_\Phi$ in (13.7)). The value and the properties of this unique voltage for a silicon p-n junction are presented in Section 13.1.2.1.

The core piece of a bandgap reference consists of two bipolar transistors and an op amp servo loop (Figure 13.4(a)). The circuit has more freedom degrees: (ratio of the collector currents, ratio of the emitter areas, etc.).

Considering the version with equal collector resistances $R_3 = R_4$, the servo loop and the two equal resistors force identical currents to flow through $T_1$ and $T_2$ ($I_{C1} = I_{C2}$).

The voltage drop on the resistor $R_2$ is given by the difference of the BE voltages of $T_1$ and $T_2$, which is still dependent on the transistors emitter areas ratio ($n = A_{E2}/A_{E1}$):

$$I_{C2}R_2 = V_{BE1} - V_{BE2} = \frac{k \cdot T}{q} \cdot \ln \frac{I_{C1} \cdot A_{E2} \cdot I_{S2}}{I_{C2} \cdot A_{E1} \cdot I_{S1}} = \frac{k \cdot T}{q} \cdot \ln \frac{A_{E2}}{A_{E1}} \tag{13.22}$$

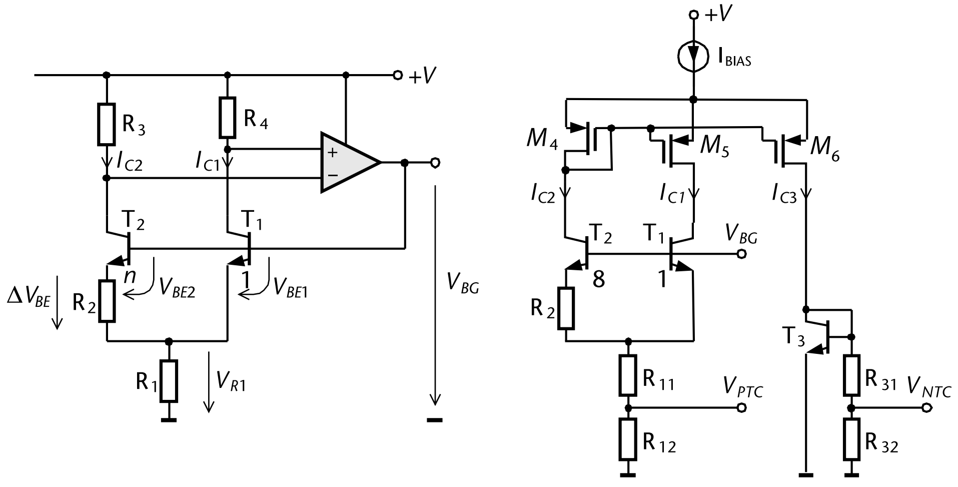
and, hence, the collector currents are given by:

$$I_{C2} = \frac{1}{R_2} \frac{k \cdot T}{q} \cdot \ln n = I_{C1} \tag{13.23}$$

**Table 13.1** Overview of Commercial Integrated Temperature Sensors

| *Part Number* | *Output Type* | *Temperature Range* | *Output Characteristic* | *Accuracy (Typical)* |
|---|---|---|---|---|
| LM34 | Analog | −50 to +300°F | $V_{OUT} = 10$ mV/°F · $T^{[°F]}$ | ±1°F |
| LM35 | Analog | −55 to +150°C | $V_{OUT} = 10$ mV/°C · $T^{[°C]}$ | ±0.5°C |
| LM45 | Analog | −20 to +100°C | $V_{OUT} = 10$ mV/°C · $T^{[°C]}$ | ±3°C |
| LM50 | Analog | −40 to +125°C | $V_{OUT} = 500$ mV + 10 mV/°C · $T^{[°C]}$ | ±2°C |
| LM60 | Analog | −40 to +125°C | $V_{OUT} = 424$ mV + 6.25 mV/°C · $T^{[°C]}$ | ±2°C |
| LM75 | Digital, I2C | −55 to +125°C | 9-bit digital temperature readings from an integrated Sigma-Delta ADC | ±3°C |
| LM135 | Analog | −55 to +150°C | $V_{OUT} = 10$ mV/°K · $T^{°K}$ | ±1°K |

*From:* [9].



(a) (b)

**Figure 13.4** Bandgap references with (a) op amp's servo loop, and (b) current mirror.

The currents flow together through R$_1$ and generate a voltage:

$$V_{R1} = R_1 \cdot \left(I_{C1} + I_{C2}\right) = 2\frac{R_1}{R_2}\frac{k \cdot T}{q} \cdot \ln n \tag{13.24}$$

Obviously, this voltage is a PTAT voltage. The generation of a PTAT voltage represents a first feature of the circuit.

This voltage adds up to $V_{BE1}$ and the sum could be—under certain circumstances—temperature-independent. Because $V_{BE1}$ has a temperature coefficient of −2 mV/°K, the voltage $V_{R1}$ has to have a temperature coefficient of +2 mV/°K. The derivative of $V_{R1}$ provides an expression:

$$\frac{dV_{R1}}{dT} = \frac{2R_1}{R_2}\frac{k}{q} \cdot \ln n = +2\ \frac{\text{mV}}{°\text{K}} \tag{13.25}$$

which permits to determine the design condition to meet the requirement +2 mV/°K.

For usual values $n = 8$ ($\ln 8 \approx 2$), the resistors ratio becomes:

$$\frac{R_1}{R_2} = \frac{2\ \text{mV/°K}}{2 \cdot \frac{k}{q} \cdot \ln n} = \frac{2\ \text{mV/°K}}{4 \cdot 86.17 \cdot 10^{-3}\ \text{mV/°K}} \approx 5.8 \tag{13.26}$$

The output voltage $V_{BG}$ is the sum of the two voltages:

$$V_{BG} = V_{BE1} + V_{R1} = V_{BE1} + \frac{2R_1}{R_2} \cdot V_T \cdot \ln n \tag{13.27}$$

If the demand (13.25) is met, this equation can be inserted in (13.7). Hence, (13.7) becomes:

$$V_\Phi = V_{BE1} + \frac{2R_1}{R_2}\frac{k}{q} \cdot \ln n \cdot T \tag{13.28}$$

The comparison of (13.27) and (13.28) leads to the final formula:

$$V_{BG} = V_\Phi \tag{13.29}$$

which explains the name of bandgap reference.

In summary, the voltage $V_{BE1}$ is complementary to the absolute temperature, and summing it with a properly proportioned $V_{PTAT}$ from across $R_1$ ($V_{R1}$) gives the desired end result, where the output voltage $V_{BG}$ becomes constant with respect to temperature and has a value of 1.25V (typical).

The generation of a temperature-independent bandgap voltage represents the second feature of the virtues of this circuit in Figure 13.4(a), which is popularly known as the Brocaw Cell after its inventor [10].

Apart from schematics with op amp servo loops, the ASICs usually contain bandgap cells on transistor level (Figure 13.4(b)). The circuit is very similar to those in Figure 13.4(a). The significant change is the replacement of the servo loop by a current mirror $T_4 - T_5$, which equalizes the collector currents $I_{C1} = I_{C2}$. The resistor $R_1$ is realized as a voltage divider to give the freedom in the fixing of the PTAT-based voltage $V_{PTC}$, which has the same positive temperature coefficient.

It is advisable to locally generate a complimentary voltage $V_{NTC}$ with a negative temperature coefficient. This could be produced using the Figure 13.2(b) cell, which offers high performances, but also using the easy Figure 13.2(a) as a source. This is represented in Figure 13.4(b) by the transistor $T_3$ that is supplied with a current $I_{C3} = I_{C1}$. For the same flexibility reasons, the high impedance resistor $R_3$ is here again divided.

Bandgap references are widespread in electronics. There is almost no integrated circuit without an internal bandgap. The device is mentioned many times throughout this book.

There is a large variety of discrete or integrated, bipolar or CMOS references containing only the bandgap core as well as additional stages, such as multipliers of the value $V_{BG}$, output buffers, starter circuits, and temperature drift curvature correction stages. The number of granted patents disclosing bandgap improvements is also impressive.

High-precision references, which are calibrated by the manufacturer, are frequently used (e.g., LM4040 [11]). The almost zero temperature coefficient of the reference voltage can be fine-tuned by the resistor $R_2$ (see (13.23)) and its value at +27°C (300°K) by $R_1$ (Figure 13.4). A binary 2-bit adjustment and the storage of the trim values by microfuses blowing or in nonvolatile memories is currently the standard procedure. This trimming normally occurs during the IC probing.

## 13.2 Theoretical Considerations Regarding the Temperature Behavior of ISEs

To develop a model for the temperature behavior of inductive sensing elements, we can start with the evaluation of experimental measurements made with the resonant

LC-circuit mounted as a standard sensing element (see Figure 3.20 and the Maxwell model in Figure 4.16) and connected to the test pilot oscillator (Section 8.3.2, Figure 8.17). Frequency sweeps of the Q factor of the resonant LC circuit were performed under the following two conditions:

1. Without a target in front of the sensing element, at $T = +25°\text{C}$ (Figure 13.5(a));
2. With the standard target maintained at a fixed distance $s$ = 3 mm and for three ambient temperatures: $T = -25°\text{C}$, $T = +25°\text{C}$ and $T = +75°\text{C}$ (Figure 13.5(b)).

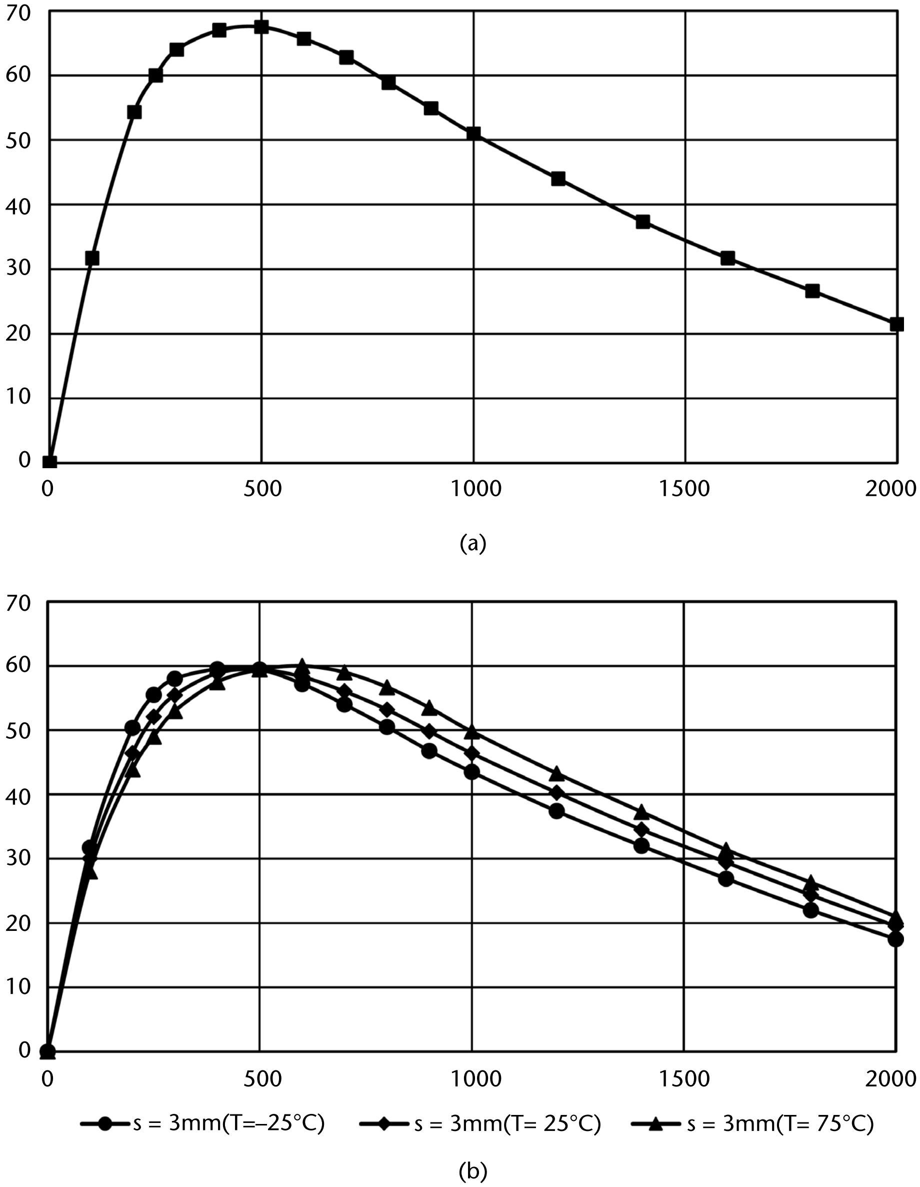


**Figure 13.5** Experimental results: Frequency sweep of the Q factor: (a) without target, and (b) with target at $s$ = 3 mm and at three temperatures. The frequency unit on the $x$-axis is kHz.

The frequency sweep in Figure 13.5(a) evidences two ranges:

1. *Lower frequency range:* $f \leq 500$ kHz for the tested sample. The behavior of the Q factor is primarily determined by the losses in the coil winding and corresponds to the incipient straight-line segment with the equation:

$$Q_W = \frac{1}{D_W} \approx \frac{\omega L}{R_W} \tag{13.30}$$

   where $\omega = 2\pi f$ is the angle frequency, $L$ and $R_W$ are the winding inductance and resistance, and $D_W$ is the corresponding loss factor;
2. *Higher frequency range:* $f > 500$ kHz. The behavior of the Q factor is especially determined by the losses in the coil core represented by a loss resistance $R_C$. The curve shape can be sufficiently accurately represented by a hyperbolic function such as:

$$Q_C = \frac{1}{D_C} \approx \frac{R_C}{\omega L} \tag{13.31}$$

The total loss, obtained by adding the partial components leads to the final expression of the Q factor:

$$Q = \frac{1}{D_W + D_C} = \frac{1}{\dfrac{1}{Q_W} + \dfrac{1}{Q_C}} \tag{13.32}$$

The maximum of the Q factor in the frequency sweep corresponds to the equality of the loss factors, which appears at the frequency:

$$\omega_{\max} \approx \frac{\sqrt{R_W \times R_C}}{L} \tag{13.33}$$

In order to analytically model the temperature behavior under the circumstances stated above, some influences can be neglected.

Thus, if the temperature factor of the core ferrite is very low $\alpha_F < 10^{-5}$ (see Table 6.1) then $L$ can vary only due to geometry changes. The thermal dilatation coefficient of the winding wire (copper) is also small ($\beta \leq 1.6 \cdot 10^{-5}$/°K). On this basis, the variation of the inductance over temperature can be neglected in comparison to the strong changes of the $R_W$ ($\alpha_{\text{Copper}} = 4 \cdot 10^{-3}$/°K).

With this simplifying assumption, the temperature dependence of $Q_W$ can be expressed by:

$$Q_W(\omega, T) = \frac{\omega L}{R_{W0} \cdot (1 + \alpha \cdot T)} \tag{13.34}$$

where $R_{W0}$ is the loss resistance at the zero absolute temperature and $T$ is the absolute temperature.

Based on the relationship in (13.33), the $Q_C$ dependence results in:

$$Q_C(\omega,T) = \frac{\omega L \cdot (1+\alpha \cdot T)}{R_{W0} \cdot \left(\omega/\omega_{\max}\right)^2} \tag{13.35}$$

Making substitutions of (13.34) and (13.35) in (13.32), the Q factor can be expressed by:

$$Q(\omega,T) = \frac{\omega L}{R_{W0}\left[(1+\alpha \cdot T) + \left(\omega/\omega_{\max}\right)^2 \cdot 1/(1+\alpha T)\right]} \tag{13.36}$$

Applying a Taylor series representation of the factor $1/(1 + \alpha \cdot T)$ [12] and neglecting factors of order $n \geq 3$ ($\alpha T << 1$), (13.36) leads to a final expression of the Q factor versus frequency and temperature:

$$Q(\omega,T) = \frac{\omega L}{R_{W0}\left\{1 + \left(\omega/\omega_{\max}\right)^2 + \alpha \cdot T\left[1 - \left(\omega/\omega_{\max}\right)^2 \cdot (1-\alpha \cdot T)\right]\right\}} \tag{13.37}$$

The temperature behavior is determined by the second term of the denominator, which is dependent on the temperature but also on the frequency:

$$F_Q = \alpha \cdot T\left[1 - \left(\frac{\omega}{\omega_{\max}}\right)^2 \cdot (1-\alpha \cdot T)\right] \tag{13.38}$$

The temperature behavior can be now better understood, dividing the frequency range into the subranges defined above:

- In the low frequency range: $\omega << \omega_{\max}$, $F_Q \approx \alpha \cdot T$ and the Q factor expression:

  $$Q(\omega,T) \approx \frac{\omega L}{R_{W0} \cdot (1+\alpha \cdot T)} \tag{13.39}$$

  shows a value drop if the temperature increases (negative temperature coefficient). The slope of the incipient linear segment of the curve $Q(\omega)$ suffers a reduction.

- In the high frequency range: $\omega >> \omega_{\max}$, $F_Q$ can be approximated by:

  $$F_Q = -\alpha \cdot T \cdot \left(\frac{\omega}{\omega_{\max}}\right)^2 \cdot (1-\alpha \cdot T) \tag{13.40}$$

  and the processed expression of the Q factor:

  $$Q(\omega,T) \approx \frac{\omega L}{R_{W0} \cdot \left(\omega/\omega_{\max}\right)^2 \cdot (1-\alpha \cdot T)} \tag{13.41}$$

shows a value growth if the temperature increases (positive temperature coefficient). The value of this coefficient is reversely dependent on the frequency.

- The limit for $\omega = \omega_{\max}$, $F_Q \approx -(\alpha \cdot T)^2 \approx 0$ and the peak of the Q factor is theoretically independent of the temperature:

$$Q\left(\omega_{\max}\right) \approx \frac{\omega L}{2 \cdot R_{W0}} \tag{13.42}$$

which is a fact that is confirmed in practice.

The analytical model matches the measurement results shown in Figure 13.5(b) and indicates two different temperature behaviors of the Q factor, with two different signs of its temperature coefficient.

The representation shows the measured frequency sweeps of the Q factor for $s = 3$ mm and at three temperatures, which are standard for such measurements. The existence of an intersection point at the frequency $\omega_{\max}$ is to be noted. By scaling the Q values for −25°C or +75°C on the Q values for room temperature (+25°C), symmetrical deviations can be observed.

## 13.3 Improvement of the Temperature Behavior by Passive Temperature Compensations

The results of the theoretical analysis in Section 13.2 lead to the specification of the ideal solution for the temperature compensation of an inductive sensor.

This should consist of an inductive sensing element exactly adjusted to work in intersection point at the frequency $\omega_{\max}$ and of a sensor electronics, which has to be designed for a minimal temperature coefficient (theoretically zero).

The professional literature on this field (publications, patents, and patent applications) discloses different solutions to improve the temperature behavior of the ISs.

Considering these sensors to be a serial data channel composed of the two mentioned units, the author has attempted to summarize the procedures found in the extensive literature in Table 13.2.

Case A includes methods based on remarkably sophisticated ISE constructions, which permit an intrinsic measuring of several parameters, including temperature, and on particular granted evaluation methods. Such realizations exhibit high performances. An example could be the EXTRA DISTANCE-Family (Series 500) provided by the sensor manufacturer Contrinex AG / Switzerland (www.contrinex.com), which contains ISs with very high switching distances, namely four times the standard values.

Case B was already introductorily defined. It refers to conventional sensing elements, which are optimally applied. Unfortunately, it is more a theoretical ideal case, very difficult to meet in reality, due to the inherent sample deviations.

Case C refers to a paradox situation and is the most preferred case. Despite a moderate temperature behavior, the method is based on a cognizable, dominant source of the system temperature drift. The shifted oscillating frequency downgrades

**Table 13.2** Overview of the Temperature Compensation Procedures for ISs

| *Case* | *ISE* | *Inductive Sensor Electronics* |
|---|---|---|
| A | Special, high-performance embodiments, designed to have minimal temperature dependence as well as local measurement and compensation possibilities. | Suitable evaluation electronics, optimized to provide high performances and good temperature stability. |
| B | Sensing element optimally applied, for example, at a working frequency that ensures minimal temperature drift of the Q factor (intersection point). | Suitable evaluation electronics, optimized to have an almost zero temperature drift. |
| C | Sensing element intentionally used at a working frequency, which differs from the optimal value and causes a predictable temperature drift of the Q factor. | Suitable evaluation electronics, optimized to compensate the temperature drift of the ISE (usually analog compensation). |
| D | Sensing element optimally applied, for example, at a working frequency that ensures minimal temperature drift of the Q factor (intersection point). | Active electronic temperature compensation techniques, which uploads adjustment coefficients or performs self-teaching adjustments. |

the sensitivity of the inductive sensing element but offers a temperature coefficient with defined sign and with a value, which deviates to a lesser extent due to the component tolerances.

The usual compensation element is the thermistor, which is connected to a network with adjustment resistors of the evaluation electronics. A classic example therefore, is the resistor $R_{ADJ}$ available at almost all oscillators in Chapter 8 (e.g., Figure 8.34(b)).

Let us consider the Q factor of the SE with a negative temperature coefficient (13.39) as an example. The oscillator output voltage will have the same behavior. Now, if the temperature rises, the value of the network $R_{ADJ}$ in parallel to the suitable NTC thermistor decreases, the feedback current in the oscillator increases and compensates for output signal reduction. This popular method has its limits: it can cover narrow temperature ranges, and the NTC has a nonlinear temperature dependence and a large sample deviation. Due to this size, it is placed in the sensor far away from the ISE and hence carries out an erroneous measurement of the temperature to be compensated.

Case D is a typical example of transferring the difficult temperature compensation task of the unfavorable sensing element into the evaluation electronics. A literature research offers a limited number of such applications, which are usually digital implementations. A concrete solution successfully implemented in serial production of high-end sensors is examined in Section 13.5.

## 13.4 Active Analog Hardware-Based Temperature Compensation Methods of ISs

The active temperature compensation techniques define two classes: one which uploads adjustment coefficients and the other that performs self-teaching adjustments.

Hence, the temperature compensation procedures could be considered similar to the linearization methods: analog hardware-based linearization (Section 9.4.1) and software-based linearization (Section 9.4.2).

Consequently, this section and the next deal with the following temperature compensation techniques:

- Analog hardware-based compensation with statistically determined batch settings;
- Software-based compensation with individual self-teaching by putting an LUT procedure into operation during normal operation.

The compensation procedure and the circuitry in the present section relies on the technical teaching description and proposed embodiment, which are claimed in the granted patent (see [13]), and exemplify the first type of the compensations mentioned above.

Performance oscillators for inductive sensors, described in Chapter 8, have an additional input to provide temperature compensation. Thus, the circuit in Figure 8.34(b) has the input SLP (slope) primarily to adjust the slope of the oscillator characteristic, that is, oscillator output signal versus distance target to sensing element. The input current can be set by the resistor $R_{SLP}$ but can be injected by an external current sink connected between this terminal and GND.

If this current sink is realized to be temperature-controlled and if the value and temperature dependence of its output current are determined and implemented, the mechanism defines a compensation technique. According to [13], the circuit that monitors the internal sensor temperature and generates the compensation current $I_{SLP}$ has a triple current sources architecture followed by a current summation of the three partial currents:

1. A rising current $I_{T1}$ over the temperature range with a positive constant temperature coefficient $TC_1$ and a value $I_{T1\_25}$ at the room temperature adjustable by an external resistor $R_{T1}$.
2. A falling current $I_{T2}$ over the temperature range with a negative constant temperature coefficient $TC_2$ and a value $I_{T2\_25}$ at the room temperature adjustable by an external resistor $R_{T2}$.

   The weighted sum of these partial currents is a linear current having room temperature values in the range [0 to $I_{T1\_25} + I_{T2\_25}$] and temperature coefficients in the range [$TC_2$ to $TC_1$]. Frequently, this current is sufficient to perform good temperature compensation.
3. For temperature ranges that are too large, the practice often shows an accentuated drift tendency of the oscillator signal at extreme temperatures (e.g., below −25°C or above +75°C). Linear compensation currents are not satisfactory; nonlinear components at very low and very high temperatures are needed. Thus, the system generates a third current $I_{T3}$, which is different from zero only in these border ranges. The slope of this dependence, and hence the additional amount, is adjustable by an additional resistor $R_{T3}$.

The general graph of the total current $I_{SLP}$ after summation (solid curve) versus temperature for a certain setting is shown in normalized representation in Figure

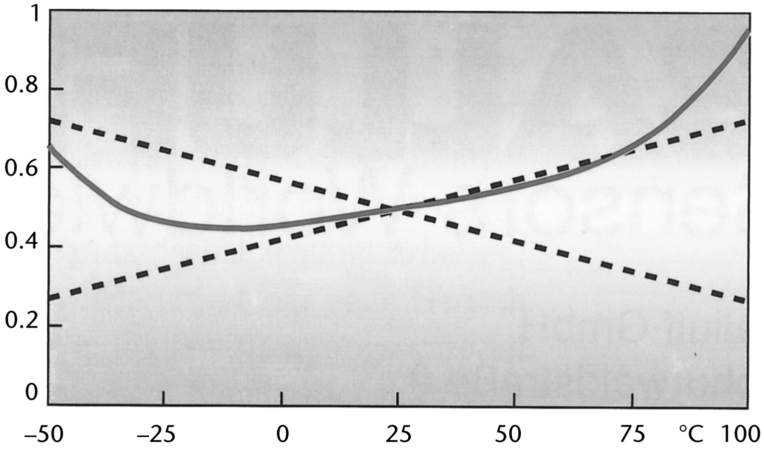


**Figure 13.6** The current summation of temperature-dependent components leads to a compensation current (solid line), whose temperature profile can be linearly and/or nonlinearly adjusted [13].

13.6. The dashed graphs outline the linear components. The influence of the addition current $I_{T3}$ is visible through the aircraft wings.

The above-defined architecture can be designed in different versions, and for bipolar or CMOS processes. Figure 13.7 shows a simple, just-enough circuitry consisting of following four subcircuits:

1. The first current generator realized as classical voltage-controlled current sink with the op amp $OA_1$ and the transistor $M_1$. It is controlled by the voltage $V_{PTC}$ provided by the multifunctional stage in Figure 13.4(b):

$$V_{PTC} = 500 \text{ mV, typ. at } 25°\text{C} \tag{13.43}$$

   and generates the partial current with positive TC:

$$I_{T1} = \frac{V_{PTC}}{R_{T1}} = \frac{500 \text{ mV}}{R_{T1}} \text{ at } 25°\text{C and } TC_{I_{T1}} = +3{,}300 \ \frac{\text{ppm}}{°\text{C}}, \text{ typ.} \tag{13.44}$$

2. The current generator realized as a second voltage-controlled current sink with the op amp $OA_2$ and the transistor $M_2$. It is controlled by the voltage $V_{NTC}$ provided by the same multifunctional stage:

$$V_{NTC} = 500 \text{ mV, typ. at } 25°\text{C} \tag{13.45}$$

   and generates the partial current with negative TC:

$$I_{T2} = \frac{V_{NTC}}{R_{T2}} = \frac{500 \text{ mV}}{R_{T2}} \text{ at } 25°\text{C and } TC_{I_{T2}} = -3{,}300 \ \frac{\text{ppm}}{°\text{C}}, \text{ typ.} \tag{13.46}$$

   To summarize, adding the adjusted currents $I_{T1}$ and $I_{T2}$, which have fixed TC values but different values at room temperature, the result is a linear

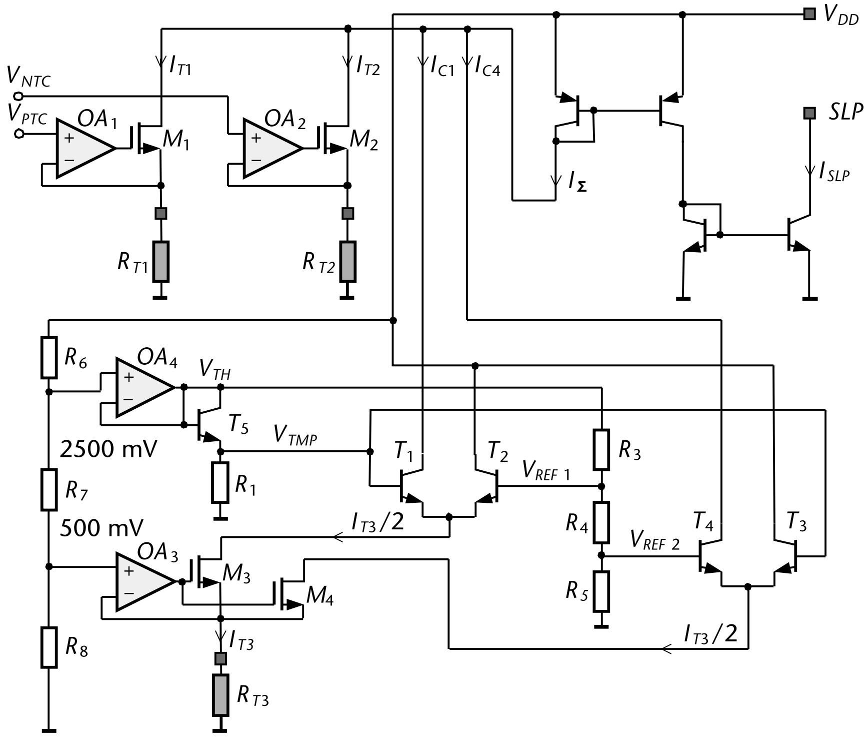
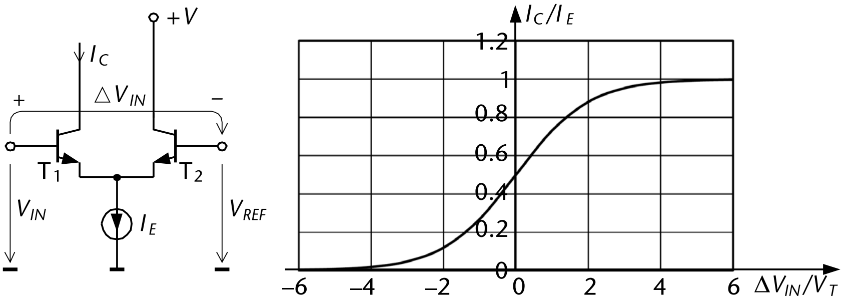

$V_{DD}$
$I_{T1}$
$I_{T2}$
$I_{C1}$
$I_{C4}$
$V_{NTC}$
$V_{PTC}$
$OA_1$
$OA_2$
$M_1$
$M_2$
SLP
$I_{SLP}$
$I_{\Sigma}$
$R_{T1}$
$R_{T2}$
$R_6$
$OA_4$
$V_{TH}$
$T_5$
$V_{TMP}$
2500 mV
$R_1$
$T_1$
$T_2$
$R_3$
$V_{REF\ 1}$
$R_7$
500 mV
$I_{T3}/2$
$R_4$
$T_4$
$T_3$
$V_{REF\ 2}$
$OA_3$
$M_3$
$M_4$
$R_5$
$R_8$
$I_{T3}$
$R_{T3}$
$I_{T3}/2$

**Figure 13.7** Schematic of the generator for the compensation current $I_{SLP}$ having a triple current sources architecture followed by a current summation. The circuit is a smart replacement of the resistor $R_{SLP}$ in Figure 8.34(b).

sum current having the needed value at room temperature and a temperature coefficient adjustable in the entire range TC ∈ [–3300 ppm/°C, +3300 ppm/°C]. For equal components at room temperature, the sum current has a double value and TC ≈ 0 ppm/°C.

3. The explanation of the current generator for $I_{T3}$ requires some additional information. There are more possibilities to implement this stage. To generate the aircraft wings, the schematic in Figure 13.7 uses two differential amplifiers and applies a property of the bipolar implementation of this amplifier



(a) (b)

**Figure 13.8** (a) Bipolar differential amplifier and (b) its transfer characteristic.

(Figure 13.8(a)). Its output, namely the collector current $I_C$ has a nonlinear dependence on the differential input voltage $\Delta V_{\mathrm{IN}}$ given by the hyperbolic tangent function [7]:

$$I_C = \frac{I_E}{2} \cdot \left(1 + \tanh\frac{1}{2} \cdot \frac{\Delta V_{\mathrm{IN}}}{V_T}\right) \tag{13.47}$$

where $I_E$ is the emitter current provided by the current source and $V_T$ is the often mentioned thermal voltage (see (8.103)).

The shape of this function shows a realistic active input span of $\Delta V_{\mathrm{IN}} \approx 8\ V_T = 208$ mV and an inflexion point with $I_C = I_E/2$ at $\Delta V_{\mathrm{IN}} = 0$ (Figure 13.8(b)).

The differential amplifier can be twice recognized in Figure 13.7, implemented with $\mathrm{T}_1$ and $\mathrm{T}_2$ and with $\mathrm{T}_3$ and $\mathrm{T}_4$. The current $I_{T3}$ generated by the current sink realized with op amp $\mathrm{OA}_3$ is adjusted with the third external resistor $\mathrm{R}_{\mathrm{T3}}$. $\mathrm{OA}_3$ is referenced by a constant voltage of 500 mV, so that its current:

$$I_{T3} = \frac{500\ \mathrm{mV}}{R_{T3}} \tag{13.48}$$

is temperature-independent. Moreover, this current is divided in half by the transistors $\mathrm{M}_3$ and $\mathrm{M}_4$ to supply the emitters of $\mathrm{T}_1$ to $\mathrm{T}_4$.

The constant reference voltages $V_{REF1} = 2000$ mV and $V_{REF2} = 1700$ mV for the differential amplifiers are generated by the voltage divider $\mathrm{R}_3 - \mathrm{R}_4 - \mathrm{R}_5$.

On the other hand, a temperature dependent voltage $V_{TMP}$ is created by a series circuit of a buffered constant voltage source $V_{TH} = 2500$ mV and the BE junction of $\mathrm{T}_5$ (see Section 13.1.2.1):

$$V_{TMP} = 2{,}500\ \mathrm{mV} - V_{BE5} = 1{,}850\ \mathrm{mV} - 2\ \frac{\mathrm{mV}}{^\circ\mathrm{C}} \cdot \left(T^{[^\circ\mathrm{C}]} - 25\right) \tag{13.49}$$

After this description of the third compensation source for very low/very high temperatures, the operation can be followed by means of the graphs in Figure 13.9. Note that the current curves (tanh functions) are represented in a simplified manner by straight-line segments.

Corresponding to the chosen references $V_{REF1}$ and $V_{REF2}$, the inflection points appear at −50°C and +100°C. The spans of the segments are ≈ 200 mV/2 mV = 100°C. Hence, the currents gently begin to manifest below 0°C and above +50°C.

4. Finally, the four currents generated by the current sinks are added in a busbar and the total current $I_\Sigma$ is twice mirrored (up and down) to get the current sink $I_{SLP}$ having an output with lower saturation level.

   This circuit has been integrated in high-volume ICs for different levels of accuracy and performs a satisfactory temperature compensation. The minimized circuit version in Figure 13.7 is suitable for additional improvements

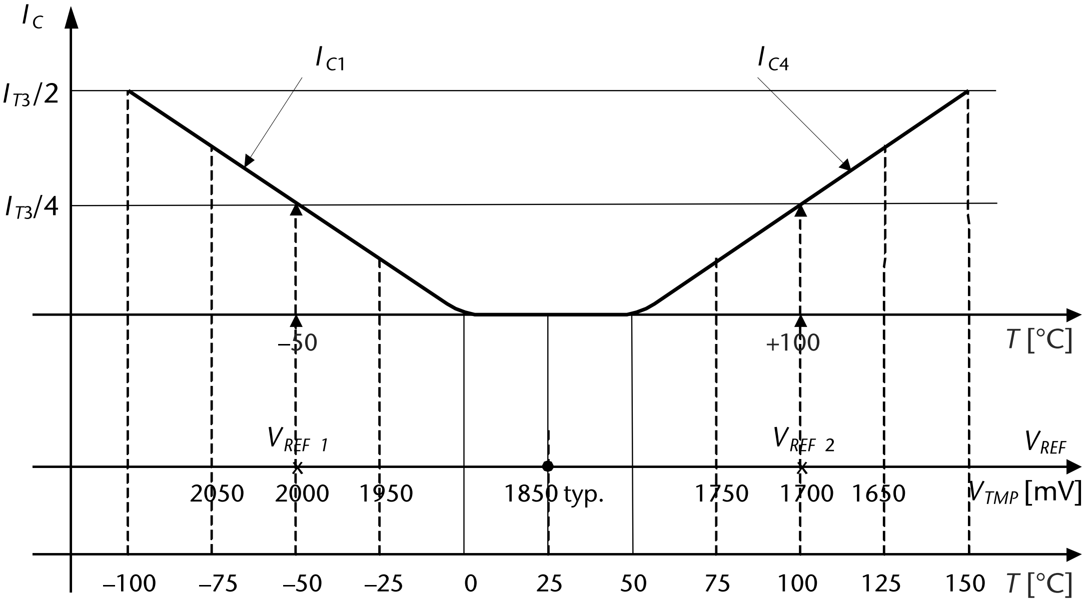


**Figure 13.9** Graphical representation of the compensation currents $I_{C1}$ and $I_{C4}$ over a large temperature range.

and adequate for bipolar or CMOS integrations. Certainly, a minimal number of five bipolar transistors (see $T_1$ to $T_5$) is a demand for all processes.

## 13.5 Active Digital Temperature Compensation Methods of ISs

The permanent improvements of the ISE and evaluation electronics makes possible to detect great distances with a safety margin (signal change between cases with the target at the desired distance and without target) of above 1% [14]. Compared to inductive sensors for standard rated operating distances (Table 2.3), which operates with a safety margin around 20%, a headroom of only 1% is very difficult to manage over wide temperature ranges, whereby the temperature drift of the Q factor is often larger ($\Delta Q_T \approx 3\%$) even for partially temperature compensated sensors.

A proven strategy to get inductive sensors that meet the hard demand of 1% headroom is to perform an evolutive approach consisting of three steps [14].

First, the ISE has to be constructively optimized: narrow mechanical and electrical tolerances of the components (magnetic core, coil wire, coil former, etc.), best quality, and small tolerances of the resonant capacitor. The key factor herein is an optimal setting of working frequency (see Case B in Table 13.2).

Second, the changes that occur over the entire temperature range and lead to a deviation to the optimal frequency corresponding to the intersection point have to be caught by an active temperature compensation. A circuit similar to those in Figure 13.7 can be monolithically integrated with the sensor electronics to get an efficient analog stage, which widely compensates the sensor for temperature influences.

Third, to manage headrooms close to 1% it is wise to apply a high-level compensation mechanism acting in the range of comparator. The generic schematic in

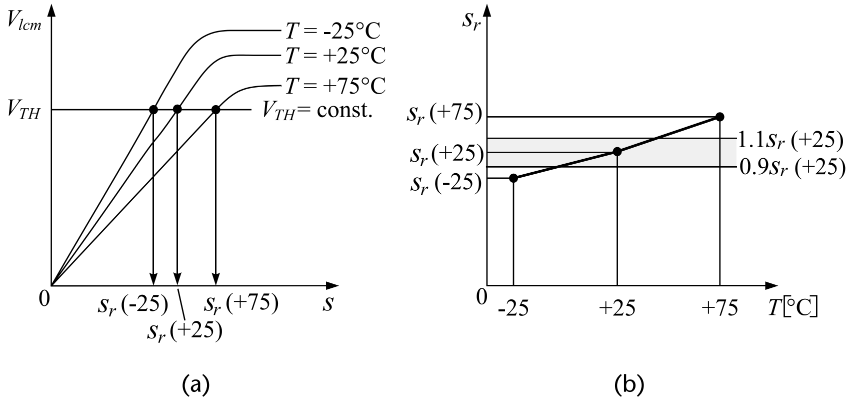


**Figure 13.10** (a) Oscillator characteristic $V_{lcm}$ versus distance $s$ at different temperatures, and (b) the resulting temperature drift of the real switching distance $s_r$ if the comparator threshold is constant.

Figure 7.1 illustrates the signal path from the ISE to such a comparator CMP. After the aforementioned wide temperature compensations and a detection in the stage DET, the signal reaches the comparator, which compares it with a normally constant threshold voltage $V_{TH}$ (Figure 13.10(a)). The still-present minimal oscillator temperature drift (exaggeratedly represented in Figure 13.10(a) by rotations of the oscillator characteristic) causes different intersections points with the horizontal line $V_{TH}$ = const. and, thus, a translation to higher values of the switching distance. In extremis, this temperature drift could exceed the standard tolerance range (gray shaded range in Figure 13.10(b)).

The improved procedure is illustrated in Figure 13.11(a). The previously mentioned digital high-level compensation now acts on the threshold voltage $V_{TH}$ and makes it temperature-dependent, so that the intersection points have the same abscissa $s_r$ (+25). The real switching distance versus temperature has now a zigzag aspect (Figure 13.11(b)), but the evolution remains within the tolerance window.

This is essentially the mechanism of the third level of the digital compensation. The internal temperature, which is measured in the coil of the ISE controls

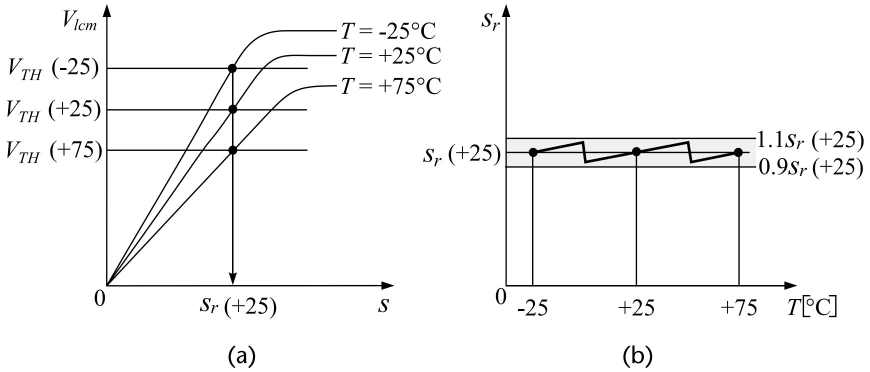
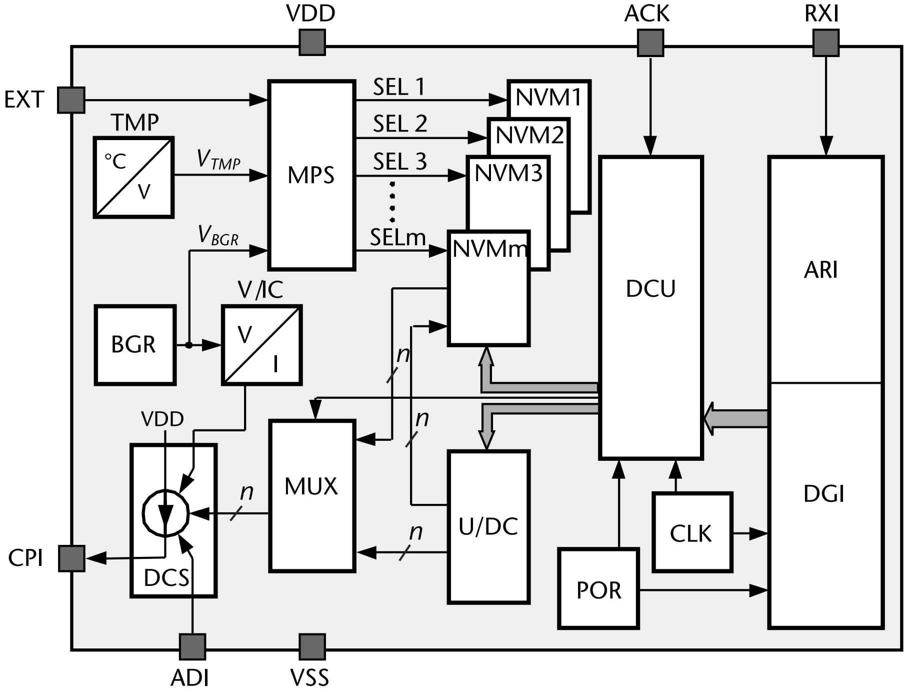


**Figure 13.11** (a) Oscillator characteristic $V_{lcm}$ versus distance $s$ at different temperatures and temperature-correlated thresholds, and (b) the resulting temperature drift of the real switching distance $s_r$.

the comparator threshold. After the final sensor assembly, individual or batch correction values are stored in a sensor assignment table. The communication for it occurs ditto over the supply lines (see Chapter 12). During the normal operation, the internal temperature sensor addresses the corresponding location of this allocation table; the correction value is extracted and correlates the threshold with the current temperature.

The separation of the intervention locations: oscillator for trimming (Chapter 12) and comparator for temperature compensation enable the parallel use of both procedures and devices.

A generic functional diagram for it is represented in Figure 13.12 [15]. There are many similarities with the generic functional diagram for the electronic sensor



List of acronyms and initialisms:

**ACU: Actuating unit**
DCS: Digital controlled current source
**TMU: Temperature management unit**
TMP: High precision temperature sensor
MPS: Memory page selector
V/IC: Voltage-current converter
BGR: Band-gap reference

**CU/P: Control unit and peripherals**
U/DC: Up/down counter
NVM: Non-volatile memory
MUX: Digital multiplexor
DCU: Digital control unit
CLK: Clock generator
POR: Power-on reset

**C/PU: Communications and programming unit**
ARI: Analog receiving interface
DGI: Digital interface

**Figure 13.12** Generic functional diagram of ASICs for temperature compensation of binary inductive sensors.

trimming ASIC (Figure 12.7). In addition, the integration process and the electrical features are identical. For this reason, the description of the units CU/P and C/PU can be taken over from Section 12.2.2.

The two new specific units are

1. Actuating unit (ACU), which is the interface of the compensation ASIC with the electronics of the sensor to be compensated for temperature and contains the stage:
   a. Digital controlled current source (DCS). Conceptually, this stage is a digital-to-analog current convertor, which sources a current flowing from $V_{DD}$ to and having the time evolution in form of an incremented or decremented digital current ramp.
2. Temperature management unit (TMU), which permanently measures the temperature and addresses the allocation table. The unit is composed of the following electronic stages:
   a. High-precision temperature sensor (TMP);
   b. Memory page selector (MPS), which performs a temperature dependent enabling of the corresponding NVM page with already stored correction value;
   c. Voltage-current convertor (V/IC) and bandgap reference (BGR) that are auxiliary stages for the DCS and MPS.

Specifically for generic functional diagrams, the block could have different realization forms.

An embodiment for the DCS is shown in Figure 13.13. The reference voltage $V_{BGR}$ = 1200 mV is converted into current $I_{REF}$ = 12 $\mu$A by the internal resistor $\mathrm{R_{REF}}$, divided by 4 in the current mirror $\mathrm{CM_1}$ and finally injected in the series ladder $\mathrm{R_0}$ to $\mathrm{R_4}$. Its resistors are binary weighted. They result in making series subcircuits using the same optimized elementary resistor, which has the resistance of $R_E$ = 3.2

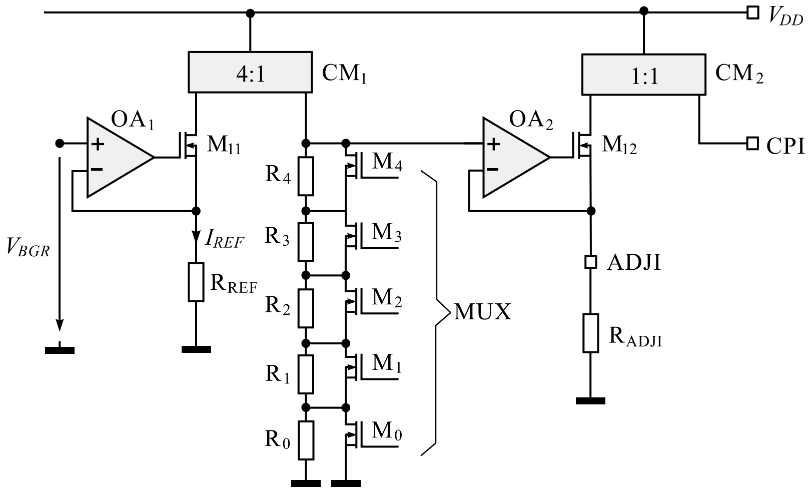

$V_{DD}$
4:1
$\mathrm{CM_1}$
1:1
$\mathrm{CM_2}$
$\mathrm{OA_1}$
$\mathrm{M_{11}}$
$\mathrm{OA_2}$
$\mathrm{M_{12}}$
CPI
+
−
+
−
$\mathrm{R_4}$
$\mathrm{M_4}$
$\mathrm{R_3}$
$\mathrm{M_3}$
$\mathrm{R_2}$
$\mathrm{M_2}$
$\mathrm{R_1}$
$\mathrm{M_1}$
$\mathrm{R_0}$
$\mathrm{M_0}$
MUX
$V_{BGR}$
$I_{REF}$
$\mathrm{R_{REF}}$
ADJI
$\mathrm{R_{ADJI}}$

**Figure 13.13** Circuit diagram of an embodiment for the stage DCS in Figure 13.12.

kΩ and is almost temperature-independent (Section 12.2.2.1). Hence, the ladder resistances varies from $R_0 = R_E = 3.2$ kΩ to $R_4 = 16 \cdot R_E = 51.2$ kΩ. The switches $M_0$ to $M_4$ are identical, every one consisting of an elementary switching transistor (Figure 12.9).

The voltage digital 5-bit ramp, which runs between 0 and 297.6 mV controls the main current source with $OA_2$, which is adjusted by an external resistor connected to the pin ADJI. The value of this resistor $R_{AJDI}$ sets the upper limit of the compensation current $I_{CPI\_\max}$ delivered by the stage, and hence, the resolution of the conversion, according to the formula:

$$R_{ADJI} = \left(2^5 - 1\right) \cdot R_E \cdot \frac{3\ \mu A}{I_{CPI\_\max}} \tag{13.50}$$

In the end, the current is mirrored by the current mirror $CM_2$ to get the final current source output CPI (Figure 13.12).

*Example 13.2*

The generation of the auxiliary voltages: reference $V_{BGR}$ and temperature-voltage $V_{TMP}$ (Figure 13.12) can be also made in different ways and with several processes. The only request is the availability of some bipolar transistors.

Figure 13.14 shows a bipolar realization containing the following core pieces:

- A PTAT cell (Figure 13.3(a)) with the resistor $R_P$ and transistors $T_5$ and $T_6$, which provides a current $I_P$ with positive temperature coefficient:

$$I_P = \frac{1}{R_P} \cdot \frac{k \cdot T}{q} \cdot \ln N \text{ and } TC\_I_P = \frac{1}{R_P} \cdot \frac{k}{q} \cdot \ln N \tag{13.51}$$

  For area ratio of $N = 48$ and $R_P = 40$ kΩ results: $I_P = 2.5\ \mu A$ at 300°K and $TC\_I_P = +8.34 \cdot 10^{-3} \mu A/°K$;
- A $V_{BE}$ cell (compare with Figure 13.2(b)) with the resistor $R_N$ and the transistor $T_{11}$, which provides a current $I_N$ with negative temperature coefficient:

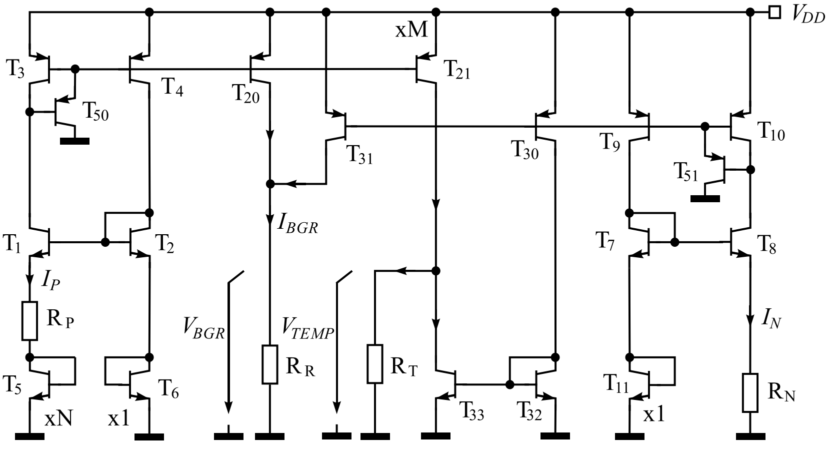


**Figure 13.14** Circuit diagram of the voltage reference $V_{BGR}$ and temperature sensor $V_{TEMP}$.

$$I_N = \frac{V_\Phi}{R_N} \cdot \left(1 - \frac{T}{2 \cdot T^*}\right) \text{ and } TC\_I_N = -\frac{V_\Phi}{2 \cdot T^* \cdot R_N} \tag{13.52}$$

For $V_\Phi = 1270$ mV, $T^* = 300°$K, and $R_N = 254$ kΩ results: $I_N = 2.5\ \mu$A at 300°K and $TC_I_N = -8.33 \cdot 10^{-3}\ \mu$A/°K;

The generation of the reference voltage $V_{BGR}$ occurs by summing the mirrored current $I_P$ (transistor $T_{20}$) with the mirrored current $I_N$ (transistor $T_{31}$). To get $V_{BGR} = 1270$ mV, the resistor $R_R$ must be:

$$R_R = \frac{1{,}270 \text{ mV}}{5\ \mu\text{A}} = 254 \text{ k}\Omega \tag{13.53}$$

For the design of the circuit for the linear temperature dependent voltage $V_{TEMP}$ the first task is to specify this linear function. The commercial part LM50 (Table 13.1) has a suitable temperature dependence expressed by:

$$V_{\text{OUT}} = 500 \text{ mV} + 10\ \frac{\text{mV}}{°\text{C}} \cdot T^{[°\text{C}]} \tag{13.54}$$

To implement this equation, the current $I_N$ (transistor $T_{33}$) is subtracted from a current derived from $I_P$ but multiplied with a factor $M$, which is the area of the transistor $T_{21}$. This resulting current flows through the resistor $R_T$ and generates the voltage $V_{TEMP}$:

$$V_{\text{TEMP}} = R_T \cdot \left(M \cdot I_P - I_N\right) \tag{13.55}$$

This equation and its derivative with respect to $T$:

$$\frac{\mathrm{d}V_{\text{TEMP}}}{\mathrm{d}T} = R_T \cdot \left(M \frac{\mathrm{d}_P}{\mathrm{d}T} - \frac{\mathrm{d}_N}{\mathrm{d}T}\right) \tag{13.56}$$

form a system of two equations, which allows to compute the unknowns $M$ and $R_T$ to reproduce (13.54): $M = 1.69$ and $R_T = 446$ kΩ.

The ASIC is also designed to work with an external temperature sensors connected to the input EXT and having the LM50 characteristic.

The temperature sensor (internal/external) controls the memory page selector MPS, which is an elementary ADC. It consists of a cascade of four windows discriminators (Section 9.5) which divide the entire temperature range in a defined number of subranges (e.g., $m - 1 = 4$ subranges). Hence, there are $m - 1 = 4$ switching temperatures usually placed in the middle of every subrange and corresponding to $m - 1 = 4$ thresholds for the MPS. The number of programming temperatures is $m = 5$ and these are identical to the limits of temperature subranges.

Correspondingly, the number of nonvolatile memory pages is $m = 5$ and according to the 5-bit resolution of the stage DCS, the locations have a capacity of 5 bit. The chosen minimal memory consumption (5 × 5 bit) and the adjustable current LSB = 0.1 $\mu$A to 1 $\mu$A were confirmed in practice as sufficiently precise.

After the presentation of the hardware aspects, the ASIC operation can be summarized by the following two tasks:

1. Every newly manufactured sensor has to be taught at more teaching temperatures (see the description above). Therefore, the sensor is placed in a trimming unit with a fixed distance to the standard target, which is equal to the wanted switching distance. A teach-in is started, the compensation ASIC changes the threshold by means of the current $I_{CPI}$ until the sensor output switches. In the end, the value of this needed current compensation value is digitally stored in the corresponding memory page, which is enabled by the temperature sensor.

   The flowchart is very similar to those in Figure 12.10. The ASIC actuating unit ACU is now different and the operation has to be performed several times. Apart from that, there are not any differences.
2. During the use of the already programmed sensor, the operation is fully automated. The temperature sensor permanently follows the temperature and enables the right memory page. The correction value is extracted and permanently controls the source DCS, while any significant temperature change occurs. In this case, the correction value is refreshed.

Section 12.2.2.4 deals with a concrete, large-scale used, high data protection communication method that has been successful and widely used for 20 years [16]. This method is also used for compensating ASIC.

The features (1 to 8) in Section 12.2 are valid. After an initial time period, it was possible to statistically process databases with correction value and to establish batch average values. In order to save time and money, the manufacturer waived the programming at five temperatures and directly loaded the batch values. There were no significant changes in the correction quality.

The ASIC in Figure 13.12 can be also used for the temperature compensation of analog inductive sensors. The output CPI is compatible with the oscillator input SLP and can compensate the oscillator for temperature variations.

Due to up-to-date sensor production technology and to application of modern ASICs, inductive binary sensors with multiple switching distances relative to the standard value are available on the market.

Just a few manufacturers can produce sensors with four-times switching distance. Balluff GmbH & Inc. (www.balluff.com) provides M12, M8 and D6.5 sensor types [17], which operate based on the above described technique. In addition to applications with limited space and highly required switching distances, such sensors (named HyperProx by Balluff) are considered by the users as real problem solvers because they have some specific features:

- They offer an equivalent double switching distance of the next larger sensor size ($4 \cdot s_n$ at M12 is equivalent to $2 \cdot s_n$ at M18);
- Despite the known dramatical reduction of the switching distance for nonferromagnetic targets (down to $\approx$ 35%), the resulting value is still higher than the switching distance of the standard version with ferromagnetic target;

- They are able to detect small metallic pieces; for example, steel balls, which move through a plastic pipe.

The aforementioned procedure can be considered a simple but efficient LUT precursor. Recent research results of patent databases provide promising real LUT applications. The granted patent [18] claims “an inductive proximity sensor having a temperature compensation circuit, which comprises a temperature sensor, ΣΔ-modulator for sampling the temperature signal and for providing a first sequence of digital values, a first digital filter stage for generating a second sequence of digital values, lookup table for receiving the second sequence of digital values and for generating a sequence of digital lookup values and a second, preferably digital filter stage with low-pass characteristic for receiving the sequence of digital lookup values and for generating a sequence of temperature compensation values, wherein the evaluation unit for receiving the result of temperature compensation values is adapted to perform a temperature compensation of the output signal on the basis of the received result of temperature compensation values.”

In conclusion, the massive spread of low-cost/low-power consumption microsystems pushes the LUT technique not only to linearize tasks but also to compensate challenges in inductive sensors.

## References

[1] NTC/PTC Thermistors for Automotive, 2016, https://www.murata.com/~/media/webrenewal/support/library/catalog/products/thermistor/r03e.ashx.

[2] Application manual PTC Thermistor POSISTOR, 2014, https://www.murata.com.

[3] Tietze, U., C. Schenk, and E. Gamm, *Halbleiter-Schaltungstechnik*, 15th Edition, Berlin Heidelberg, Germany: Spinger/Vieweg-Verlag, 2016.

[4] Hering, E., G. Schönfelder (Hrsg.), *Sensoren in Wissenschaft und Technik—Funktionsweise und Einsatzgebiete*, 2nd Edition, Wiesbaden, Germany: Vieweg+Teubner Verlag / Springer Fachmedien GmbH, 2018.

[5] Analog Devices, *Date sheet: AD693, Loop-Powered 4–20 mA Sensor Transmitter*, 2002, www.analog.com/media/en/technical.../data-sheets/AD693.pdf.

[6] Nawrocki, W., *Measurement Systems and Sensors*, Second Edition, Norwood, MA: Artech House, 2016.

[7] Gray, P. R., Hurst, P. J., Lewis, S. H., and Meyer, R. G., *Analysis and Design of Analog Integrated Circuits*, 5th Edition, John Wiley & Sons, New York, 2009.

[8] Sze, S. M., *Semiconductor Sensors*, New York: John Wiley & Sons, 1994.

[9] *TI Analog, DSP and Semiconductor Products*, 2018, http://www.ti.com/general/docs/prod.tsp.

[10] Jung, W., *Op-Amp Application Handbook*, Analog Devices, Newnes Publisher, 2006.

[11] LM4040-N/-Q1 Precision Micropower Shunt Voltage Reference, Texas Instruments, 2016, http://www.ti.com/lit/ds/symlink/lm4040-n.pdf.

[12] Nearing, J., *Mathematical Tools for Physics*, ebook, Miami: Dover Publications, 2010.

[13] Fericean, S., et. al., High-Sensitivity Proximity Sensor and Procedure to Trim It, German patent DE 10046147C1, February 21, 2002.

[14] Jagiella, M., S. Fericean, M. Friedrich, and A. Dorneich, *Mehrstufige Temperatur-Kompensation bei induktiven Sensoren*, Elektronik Fachzeitschrift, August, 2003, www.elektroniknet.de.

[15] Fericean, S., and E. Gass, Non-Contact Proximity Switch and Method for Programming It, U. S. patent 5,818,129, issued October 6, 1998.

[16] Fericean, S., F. Kruepl, M. Fritton, and T. Reider, Electronic Component for a Sensor Apparatus, Sensor Apparatus and Method of Configuring a Sensor Apparatus, U.S. patent 8,660,806, issued February 25, 2014.

[17] Balluff, *Products and Services*, 2018, https://www.balluff.com/.

[18] Machul, O., and S. Thoss, Inductive Proximity Sensor with Temperature Compensation, German patent DE 102009052467B3, July 14, 2011.

## CHAPTER 14

# Intelligent Inductive Sensors: Networking

The explosion in the use and types of communication networks over the last several decades would need an immense publication space to be fully described and treated. This chapter represents a fourth part of the book and focuses only on information regarding the specific networking of the inductive sensors.

After a short introduction, which defines the specific communication terms and parts, the next sections deal with interfaces and networking systems, which are defined and used on the lowest sensor/actuator communication level in the process-control hierarchic systems.

Hence, these sections focus on serial data interfaces, AS interface and IO link network.

Returning to the sensor output types (see Section 1.3.7), the electrical data domain—the name of the quantity used to represent or transmit information—of the sensor outputs covers three domains:

1. Analog domain, in which the information is carried by signal amplitude (i.e., voltage, current, power or charge);
2. Time domain, in which the information is carried by time relations (period or frequency, pulse width or phase);
3. Digital domain, in which signals have only two values, and the information is carried by the number of pulses or by a coded serial or parallel word.

To show the diversity of digital domain, Table 14.1 summarizes (without exhaustive pretentions) and classifies the most used systems for industrial applications [1–3].

## 14.1 Basics of the Data Communication Systems

This section examines the main attributes of these systems: topologies, models, access methods, and others.

### 14.1.1 Communication Network Topologies

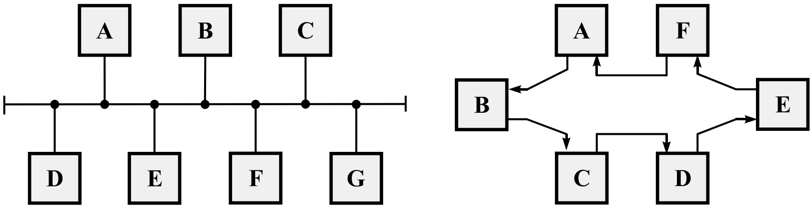
The connections between sensors, actuators, and units, which control and monitor the process, form a communication network. Its attribute topology refers to the geometrical arrangement of the network users but also to the logical configuration.

**Table 14.1** Overview of Digital Communication Systems

| *Type* | *Participants Number* | *Name of Interface* | *Master/Slave* | *Fieldbuses* |
|---|---|---|---|---|
| Serial | one to one | RS232<br>RS422<br>UART | | |
| | 1:$n$ | RS485<br>USB | | |
| | | SENSOR<br>INTERFACES | **AS Interface**<br>**IO Link** | |
| | | FIELDBUSES | Master/slave | VariNet 2<br>Bitbus<br>InterBus-S<br>SUCONET-K<br>Modnet 1<br>SINEC L2<br>Modulink |
| | | | Multimaster | PROFIBUS<br>Modnet 3<br>SINEC H2<br>LON<br>FIP<br>P-Net<br>CAN<br>DeviceNet |
| | $m$:$n$ | LAN<br>Ethernet<br>Tokenring<br>Tokenbus<br>Arcnet | | |
| Parallel | 1:1 | Centronics | | |
| | 1:$n$ | IEC 488<br>GPIB | | |
| | $m$:$n$ | VME<br>Multibus<br>SCSI | | |

Making abstraction of the network users' ranking, the following topologies are established:

- The point-to-point (PtP) topology is a simple interlink that requires separate channels (physical and logical) between every two users. The method is popular due to its simplicity and service-friendly, although it has limited communication possibilities and requires high investments (cable and connectors).
- The linear network topology contains a common physical channel for all users (A–G in Figure 14.1(a)). The hardware is considerably simpler and robust; however, the operation and access to the channel has to be rigorously defined.
- The tree network topology is an extension of the linear version suitable for large user number and/or great distances between them. The network is divided in subnetworks. Each subnetwork is controlled by a repeater, which internally manages it and outwardly realizes the link with the main tree.



(a) (b)

**Figure 14.1** The most popular communication network topologies: (a) linear, and (b) circular.

- The circular topology represents a geometrical ring layout of the PtP connections between users (A–F in Figure 14.1(b)). This involves packing the information to be transmitted in a telegram and transferring it over the bus from one user to the next one, which means that users must have a transceiver (transmitter-receiver (TCVR)) functionality. The network can cover large areas; however, it becomes critical if a user breaks down. To prevent a system shutdown the network has specific features (it can operate in both senses of rotation, can detect and isolate the defective user, etc.);
- The star topology assumes a PtP connection between every user and a central station, either being a coupling element, which couples the transmitter of a user to the receiver of another user, or having a master capability to manage the communication. The drawback is the permanent participation of the central unit; if this breaks down, the system goes unconditionally out of order.

### 14.1.2 Network Access Procedures

The operations described in Section 14.1.1 require that, at any given moment, only one transmitter be connected to the network. This implies strict access rules.

The controlled (deterministic) access methods require the transmitter to be known before the communication starts. The network allocation can be performed by a master (see the following master-slave procedure) or by a user with network coordination rights (see token procedures). The time duration can be determined and specified, which is the reason they are real-time access methods.

- The master-slave procedure is the traditional communication model and assumes the existence of a master, which establishes and coordinates the connections with subordinated devices (slaves) (Figure 14.2(a)). The slave promptly replies to a master request (immediate response). Usually, the master cyclically interrogates the slaves (polling). Nevertheless, it is possible to define priorities, thus certain slaves are called on more times during a cycle (see the direct arrows in Figure 14.2(a)).

  Usually, the hardware and software implemented in slaves is moderate; the intelligence is concentrated in the master. This fact minimizes the bus investment for a large slave number. The drawback is that an information change between two slaves exclusively occurs through the master; it takes at least two cycles (sometimes more due to additional data processing in master). If the

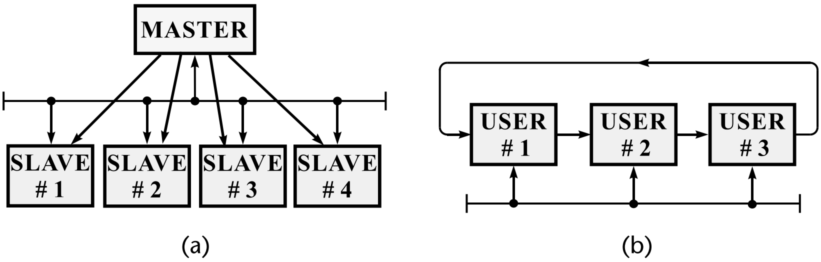


**Figure 14.2** Linear network topologies with popular access procedures: (a) master-slave, and (b) token (authorization transfer inside the logical ring).

master breaks down, the system goes unconditionally out of order and needs assistance. By contrast, if a slave breaks down, the situation is less critical; this device is removed from a participant list stored in the master, is deactivated, and ignored, waiting to be replaced.

- The token procedure, defined in the standard IEEE 802, confers equal competence to all users. The authorization for the coordination right is given by means of a particular message (token), which is in possession of every user for a defined time duration and transferred to the next one at the end of this interval. The transfer of the authorization from one user to the next occurs inside a logical ring (Figure 14.2(b)).

  This access can also be applied in a network with circular topology. The authorization rotates in the ring until it attains a user that is ready to send. It grips the authorization, switches it on busy, and attaches its message to be transmitted. The frame continues to rotate until the addressed user is achieved. The information is downloaded and the frame also rotates, having the initial transmitter as a target. This now makes a comparison between the initially sent data and the received data, and if these are consistent (no transmission errors), the authorization is released and sent to the next user.
- The token-passing procedure is a mixed method whose network contains active users, which are able to manage the communication (see the logical ring in Figure 14.3) and to change data with other active users (token principle) as well as with passive users available in the linear network (master-slave principle). The method is advantageous owing to its flexibility and speed (see PROFIBUS in Table 14.1). Every user must be able to distinguish between active and passive users.

In contrast to the deterministic access methods, systems with uncontrolled (random) access could experience a simultaneous network access demand from more users, situations that should be avoided. The arbitrary access impedes time specification of the transmission cycles. The access is regulated by several procedures:

- Carrier sense multiple access (CSMA): A user willing to send a message watches the network (carrier sense), and, if this is free, starts the transmission. If the network is busy, it cancels its intention and tries to get access later. The drawback is the collision risk, which appears when two users that watch the network begin to send quasi-simultaneously. The collision is a result of the propagation delays.

- CSMA with collision detection (CSMA/CD): This reduces the collision risk, because the transmitting user watches the line during its transmission. A difference between the sent and received data signalizes a collision. Hence, the user stops the work and sends an alert message (jam) informing the community above the collision. They step down and the disrupted user tries again to send the messages.
- CSMA with collision avoidance (CSMA/AV) is similar to the CD method but to prevent collisions, the users have priorities (extra flag in the frame address field) according to a ranking list. In case of a collision, the user with lower priority steps down to give right of way on the network. After a successful transmission of a user, its priority is switched off for a defined period of time, to avoid a continuous access of this user and access impossibility for other users.

### 14.1.3 Industrial Fieldbuses: Definition and Features

A fieldbus is defined in [6] as a digital, serial, two-way communication system, that interconnects measurement and control equipment such as sensors, actuators, and controllers. Conceptually, it provides a means to replace the PtP connectivity of users with a multidrop connection that can communicate with multiple sensors over a single communication path.

The occurrence and evolution of industrial fieldbuses is focused on two topologies (see Figure 14.1):

1. The linear bus topology (with an option for tree extensions), where a user becomes master and the remaining devices have a subordinate role (slaves). They can be coupling units to the physical bus for single sensors or they can be integrated into the sensor. The communication occurs through a single line (wire or radio) and the user position is aleatory. Advantageously, the

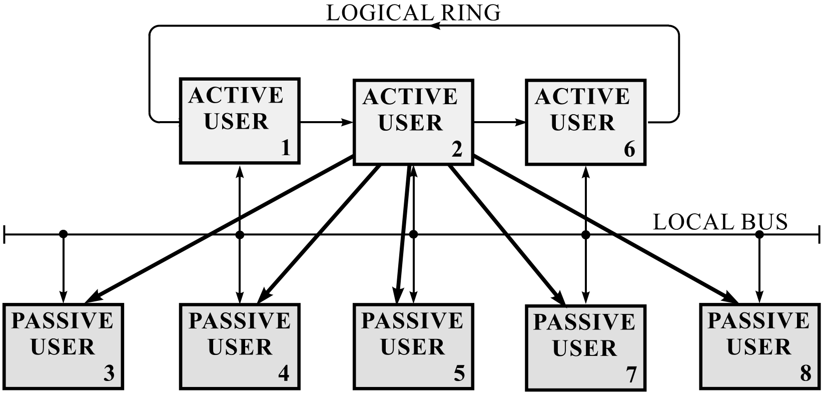


**Figure 14.3** Token-passing procedure.

system has a sole and reduced wiring hardware, the signal propagation along the bus is bidirectional, and the user number is flexible.

2. The circular bus topology consists of a regular and process-oriented arrangement of the users. Every device has two connection ports: an input and an output. There are not any hierarchic subordinations or master-slave distinction, and hence strict communication rules have to be respected. Redundancy is ensured by the implementation of data frames rotation in both senses.

Data transmission is permanently threatened by electromagnetic perturbations, which could alter the information by changing bits from 0 to 1 or vice versa. To avoid that, hardware precautions (shielding, galvanic separation, etc.), but—most important—software prevention measures (detection and correction of the corrupted information) are applied.

#### 14.1.3.1 Data Error Handling

The general methods to detect and correct damaged information can be summarized as follows:

- Use of the parity bit (even or odd). It is the easiest method and provide the lowest Hamming distance $d = 2$ [2].
- Transmission of a frame of bytes, each byte (8 bits) consisting of seven information bits followed by the eighth parity bit. The last byte in the frame contains the parity bits of the precedent bytes.
- The cyclic redundancy check (CRC) implies the computation of parity polynomial generator and the comparison of the received and local recalculated polynomial. Identity means error-free with a Hamming distance of 4 to 6 [2].

Additional prevention measures are described in the next sections, which deal with detailed network presentations.

#### 14.1.3.2 Telegram Formats

With respect to a fieldbus communication, a telegram is the coherent, directed transfer of digital characters either from master to device or from device to master.

A frame comprises a master telegram and the subsequent device telegram. A frame type represents a combination of master telegram and device telegram for the transport of certain number of process data and on-request data in both directions.

The high-level datalink (HLDL) control procedure contains a variety of protocols for a serial bit transmission, which is synchronous in start-stop running mode and has more protections.

The frame is composed of a start flag, an address byte, a control byte (expressing the frame type: data transfer, monitoring, etc.), information bytes, and the last stop byte.

The procedure states three transmission modes:

1. Normal answer mode (NRM) corresponds to the master-slave method: the slave does not spontaneously send/transmit solely at master request;

2. Asynchronous answer mode (ARM), in which spontaneous answers are allowed;
3. Asynchronous balanced mode (ABM), in which all users have identical communication possibilities.

The standard protocol Universal Asynchronous Receiver and Transmitter (UART) was initially developed for computer communications but in the meantime it became an important standard for industrial buses. The UART features in overview are

- Asynchronous data transmission in simplex, half-duplex, or full-duplex mode;
- 10 or 11-bit characters with a defined structure: start-bit (#0), seven or eight data bits (LSB up to MSB), parity bit, and stop bit.
- For time synchronization there are not any special telegrams, but the transmission speed is the same (150 bit/s to 230.4 kbit/s or more).
- Telegrams without data have a fixed length of the following six characters: start (with the command coding), receiver address, transmitter address, control, frame check sequence (FCS), and stop byte.
- Data telegrams with fixed length have a similar structure. The additional data bytes are placed before the FCS byte.
- Data telegrams with variable length could have up to 246 data bytes. They also have bytes specifying the total length and possible mode addresses (e.g., PROFIBUS has a frame with 255 bytes maximal length).

#### 14.1.3.3 Coding of the Digital Information

The physical representation of the binary values (0 and 1) is done by modifying the amplitude, frequency, phase, or signal edges of the carrier signal. Figure 14.4 and Table 14.2 summarize the conventional codes used in communication networks. Some bit patterns allow extracting of the clock frequency information. Bit patterns without DC components are more convenient as the information carrier signal can be overlaid on the supply voltage.

### 14.1.4 ISO/OSI Network Reference Model

The communication between control systems and also inside a system is facilitated if this process takes place according to previous regulations and agreements.

The ISO/OSI reference model was elaborated from the International Standards Organization and standardized in 1983 (ISO 7498). The Basic Reference Model for Open Systems Interconnection stated in ISO7498 describes the communication in open systems (i.e., systems that comply with this standard).

The model abstractly divides the communication into seven layers with well-defined functionality (Figure 14.5).

Every layer has well-specified tasks. The communication between layers through prescribed interfaces is defined, but only on an abstract level.

The user of the communication system (generally another system that needs to communicate on the network) interacts on the top layer, which is layer number 7.

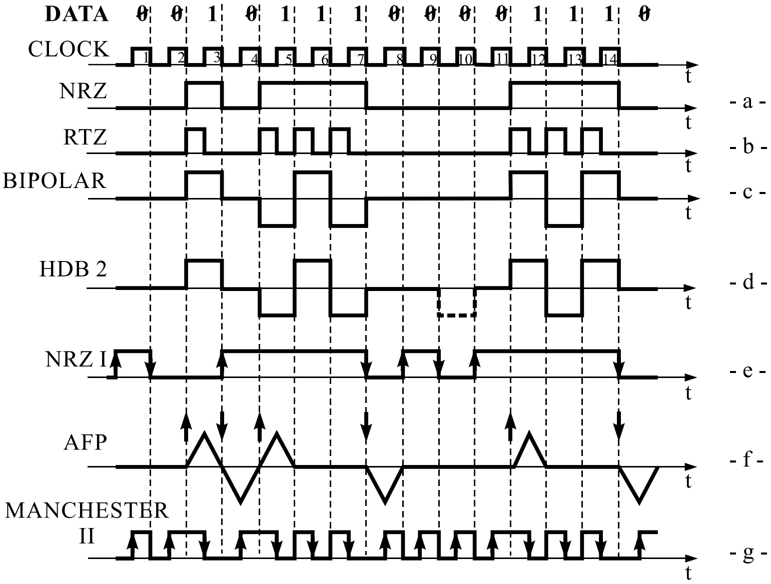

DATA
0 0 1 0 1 1 1 0 0 0 0 1 1 1 0
CLOCK
1 2 3 4 5 6 7 8 9 10 11 12 13 14
NRZ
RTZ
BIPOLAR
HDB 2
NRZ I
AFP
MANCHESTER
II
t
- a -
- b -
- c -
- d -
- e -
- f -
- g -

**Figure 14.4** Conventional bit patterns used in fieldbus implementations.

**Table 14.2** Conventional Bit Pattern Used in Fieldbus Implementations

| *Bit Pattern* | *Corresponding Figure* | *Is Clock Information Available?* | *DC Component* | *Remarks* |
|---|---|---|---|---|
| Nonreturn to zero (NRZ) | 14.4(a) | No | ≈0 | |
| Return to zero (RTZ) | 14.4(b) | No | ≈0 | |
| Bipolar | 14.4(c) | No | 0 | |
| High-density bipolar ($n = 2$) (HDB 2) | 14.4(d) | Yes (after every $n$th = second pulse) | ≈0 | After every succession of $n = 2$ bits of value 0, an additional pulse (with the polarity of the last bit of 1) is inserted |
| Nonreturn to zero insert (NRZ I) | 14.4(e) | No | ≈0 | Falling clock edge causes a polarity change for the following bit of 0 |
| Alternating flanks pulses (AFP) | 14.4(f) | No | 0 | Positive pulses for logical changes from 0 to 1 and vice versa |
| **Manchester II:** phase modulation | 14.4(g) | Yes | 0 | Changes are synchronous with the positive clock edge; namely, positive signal edge for bits of 0 and vice versa |

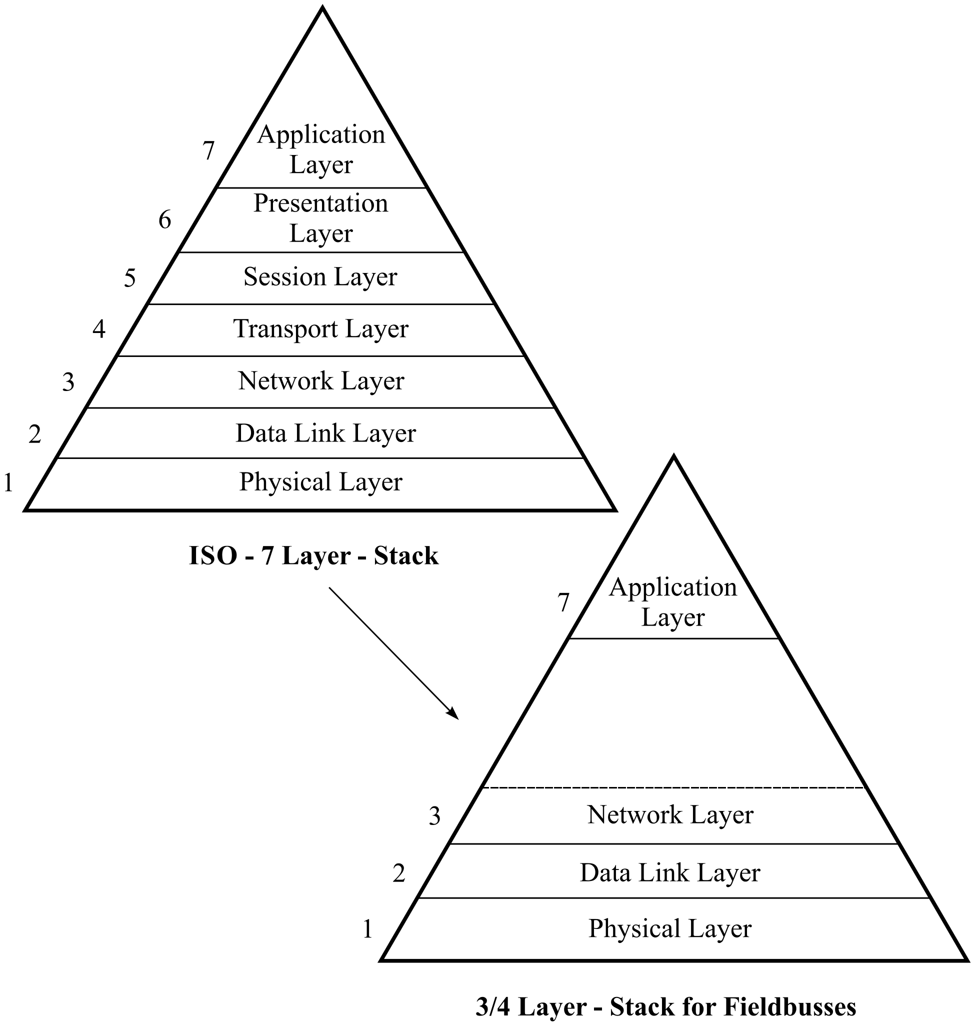

7
Application Layer
6
Presentation Layer
5
Session Layer
4
Transport Layer
3
Network Layer
2
Data Link Layer
1
Physical Layer
ISO - 7 Layer - Stack
7
Application Layer
3
Network Layer
2
Data Link Layer
1
Physical Layer
3/4 Layer - Stack for Fieldbusses

**Figure 14.5** The ISO/OSI Seven-Layer Model providing a method for segmenting communication functions (left stack) and the minimized model for fieldbus communications [3].

Most networks do not implement all of the layers in the reference model. In this case, formal definitions and features of certain layers are omitted. This is also valid for fieldbus communications. Layers 4 to 6 contain complex functions, which are not necessary in the fieldbus domain and therefore are not realized in available sensor buses.

The model decomposes an arbitrary communication network into the stack of seven layers, which are

1. Physical Layer, which is responsible for converting between the data representation of the network messages and the actual physical representation of data in the network medium. The layer specifies the behavior of the electric circuit and the physical structures of the connections. The bit coding (see

Figure 14.4) is a task on this layer. The voltage/ current levels for bits of 0 or 1, as well as the time parameters, are defined within this layer. Resuming, the physical layer manages the transmission on the bit level.

2. Data Link Layer, which is responsible for several functions. This layer manages access to the network medium, structures the bits of information into well-defined groups identified as frames or messages, handles the identification of the source and the destination stations on the network and needs to repeat a message, and provides for error-free transmission of a message from the source to destination user, all according to the data link layer protocol (see Section 14.1.3.2). The data link layer is the logic layer.
3. Network Layer, which treats the network as a logical unit and encapsulates functions related to routing of messages within a single network or among multiple networks. The layer uses addressing in a variety of forms.
4. Transport Layer, which provides any additional data transfer functions not directly stated by the data link layer for end-to-end reliable messaging. Some functions may require the use of multiple datalink layer messages to accomplish a reliable message transfer. The generation of these messages, sequential disassembly, delivery, and assembly of data is accomplished by the transport layer. It also recovers from lost, duplicated and misordered messages.
5. Session Layer, whose activity is to provide for a higher level of control of network usage and data flow, including opening, maintaining, and closing communication channels.
6. Presentation Layer, which provides functions to transform data from formats that are transportable by the network to the use-accessible formats, as defined in the application layer and understood in the local station.
7. Application Layer, which is always present in the communication model and provides communication services directly to the user application. The usage and formatting of these services is summarized in the application layer protocol. The user interacts with the network by invoking services provided by the application layer and passing data to and from the network through these services. The usual tool therefore is the File Transfer, Access and Management (FATM).

## 14.2 Requirement Description for the Sensor and Actuator Communication Level

In an up-to-date production company there is a large variety of communication networks, which are hierarchically subordinated and interconnected according to a general hierarchic model called computer integrated manufacturing (CIM). This model has five levels; however, for the purpose of this book, the lower levels are relevant; namely (a) the sensor/actuator level and (b) the production and process level.

The communication on the lowest level between sensors/actuators and the control unit (mostly programmable logic controllers (PLCs) occurs over classical wired connections or—more recently—over specific sensor buses (Figure 14.6). The demands on this level are large user numbers, short frames, and real-time behavior with cycles below 10 ms, large topology area, high EMC immunity medium,

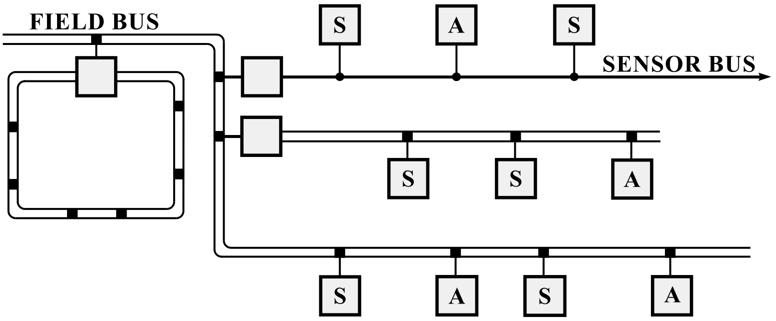


**Figure 14.6** Interconnection of the sensors (S) and actuators (A) on the sensor/actuator level and with the process management level, respectively.

reliability, low investment costs, and uncomplicated service (preferred during the running process).

The fieldbus, defined in Section 14.1.3, provides a wider range of possibilities.

The use of a fieldbus on the sensor/actuator level is conditionally appropriate. The costs are too high and the network is too laggard. However, if these features are improved, partial implementation of fieldbuses on this level makes sense. Network sensors, which are directly attachable to the bus, are ordinary sensors with network communication component added.

Finally, the interconnection to the next superordinate level usefully occurs by means of a fieldbus. The requirements here are: large operation area, flexibility and versatility, high speed, high reliability, and moderate current consumption (when the bus supplies the users by itself).

This short analysis was made to justify and promote the communication networks, which are adequate for inductive sensors and will be presented in the next sections. Their specifications describe a layered model, including the physical layer, the data link layer, and the application layer (Sections 14.6 and 14.7).

## 14.3 Intelligent Sensors

To define an intelligent sensor is not an easy task. In general, sensors are called intelligent when they can process the primarily measured values by themselves and provide improved measuring information. In addition, they can distinguish useful signals from disturbing signals and eliminate them.

The professional reference book [3] defines intelligent sensors as follows: “In an intelligent sensor, either an increase of the useful information content to the required level as a result of sensor-specific measures occurs, or the interesting benefit information is gained from the multitude of individual information, which has a low information content by itself.”

In parallel to this “information reassembly,” intelligent sensors are involved in the segmentation of the information processing. In a decentralized system, some subtasks can be locally done in subsystems containing sensors and peripheral signal processing units. Some authors consider such subassemblies to be intelligent sensors if they are spaced apart by a housing. Here a temporary data storage occurs, as well as a transfer of this information by master request or initiated by the intelligent sensor. This dialog is characterized by coding and transmission techniques, that increase the interference immunity and delimit the channel bandwidth.

Following these definitions, the need for a local signal processing unit results. In a global acceptance, an intelligent (or smart) sensor has a built-in microprocessor for automatic operation, processing data, or for achieving greater versatility. Honeywell introduced the first commercial intelligent sensor in 1983.

There is a wide variety of available intelligent sensors ranging from those that merely include digital trimming helped by an on-board microprocessor, to devices that combine data processing and communication, sometimes in a monolithic chip, for example, the ADuC7019 7029 (Analog Devices). This chip consists of a 16-bit/32-bit microcontroller with additional peripherals, and a large precision analog environment (ADCs and DACs, bandgap, comparators, oscillators and PLL, etc.) [7].

The low cost and small size of digital processors enables the enlargement of the intelligent sensor concept providing sensors with digital output capable of self-identification, self-testing, adaptive calibration, noisy data filtering, sending and receiving data, making local decisions, and so on.

These advantages are somewhat obscured by the multitude of industrial networks or fieldbuses that usually do not permit sensors to plug directly into them.

To overcome this conflict, at the end of 1990s IEEE proposed the IEEE 1451 family of Standards for Smart Transducer Interface for Sensors and Actuators with the aim of defining intelligent sensors (structure, model representation, addressing, communication protocol, etc.) to ensure sensor-to-network interoperability and interchangeability. The most important standard parts of the series are

- IEEE 1451.1, which proposes an information model for the Network Capable Application Processor (NCAP);
- IEEE 1451.2, which generally defines the transducer-to-microprocessor communication protocols as well as Transducer Electronic Data Sheet (TEDS) and the Smart Transducer Interface (STIM).

  A Transducer Independent Interface (TII) between STIM and NCAP is also proposed in form of a 10-wire serial digital interface, based on the synchronous serial communication protocol Serial Peripheral Interface (SPI) (Section 14.5).
- IEEE 1451.3, which proposes digital communication formats for Distributed Multidrop Systems (DMS).
- IEEE 1451.4, which proposes Mixed-Mode Communication Protocols (MMCP) for two-way communications of mixed-mode sensors and actuators, which connect smart sensors via conventional analog wiring.

Based on the up-to-date results, the conclusion is that the well-known sensor manufacturers used these guidelines to develop, sustain, and market standardized sensor/actuator industrial networks as Highway Addressable Remote Transducer interface (HART), AS Interface, IO Link, and so forth.

Thus, a compact and concise definition of intelligent sensors can be expressed as follows: the intelligent sensor is composed of three major functionalities: (1)

primary evaluation of the measurand and electronic signal processing, (2) microprocessor final evaluation, and (3) output information providing via a data port with semiconductor output(s). A special attribute is the existence of the microprocessor, which performs the following tasks:

- Evaluating the data from the basic sensor, possibly linearization (see Section 9.4);
- Local display;
- Providing process and diagnosis data to the programmable-logic control unit (PLC);
- Operating the communication interface;
- Performing sensor parametrization;
- Decentralized periphery (DP) for partially carrying out application software.

An adjacent question to this microprocessor refers to the parking place for it and for the sensor intelligence in relation to the decentralization and component monolithic integration. The response of [5] starts with the concept of a solid-state sensor evolution, consisting of six generations:

- First-generation devices had essentially no electronics (transducers) and the second generation contained minimal primary electronics. Electronics for signal processing were remote from the sensor.
- Third-generation devices (the state of many current sensors) already have monolithic integrated electronics (see Figure 7.2), but the hardware for network operation and the accompanying intelligence is placed into a remote unit (e.g., gateway);
- The fourth generation is a hybrid version and is realized using high-level integrated components. The communication link is two-way using a digital address and an analog high-level voltage output for data.
- The fifth/sixth-generation sensors according to [5] can be assimilated as network sensors and are able to locally provide, in a limited housing room, high-performance sensor features. The fully integrated sensor performs the data conversion itself, so that the link to the next level is a digital bidirectional bus. It is self-testing and addressable. Calibration is done via a PROM (rather than laser-trimmed resistor), which is programmed during a sensor factory test. It is also desirable to calibrate the sensor in situ on demand.

The subclassification in two generations refers to the compensation techniques: the fifth generation uploads compensation coefficients, while the sixth generation performs compensation in-module.

## 14.4 RS-232, RS-422, and RS-484 Interfaces

### 14.4.1 RS-232

RS-232 is a widely used standard method of communication and represents elements of layer 1 of the OSI model (Section 14.1.4). RS-232 is an unbalanced to ground communication only between two users (PtP). It was initially elaborated for computer communication but over time has found many industrial applications.

The interface uses up to nine communication wires and has separate pins for serial transmission of data: transmit data (TD) and receive data (RD). To distinguish the functionality of the two stations, they get different designations: data terminal equipment (DTE) or data communication equipment (DCE).

RS-232 can be used with only two signals and crossed wires: TD-pin of DTE to RD-pin of DCE and vice versa.

To increase the transmission reliability, two handshake procedures with control signals can be applied. They involve the following signals:

- RTS (ready to send) and CTS (clear to send), which inform about the request of DTE to send data and the readiness of DCE to receive data;
- DTR (data terminal ready) and DSR (data set ready), which are readiness signals of DTE and DCE to cooperate with the opposite station.

The communication steps with crossed lines: TD ↔ RD, as well as DTR ↔ DSR and with local bridged pins RTS and CTS on both sides, are:

- DTE raises the wire DTR high, indicating that data is ready to be sent and waits for the DCE confirmation. ICE receives this information at its DSR input.
- If DCE is ready to receive data, it sends its confirmation signal DTR.
- DTE receives the confirmation at the input DSR and starts the data transmission at its output TD.
- DCE receives digital data at its input RD.
- When sending is finished, DTE clears the DTR and DCE confirms by clearing its DTR.

RS-232 is typically implemented in a full-duplex fashion, since each station can transmit to the other simultaneously using separate wires. A digital 1 bit is represented by a positive voltage in the range of 5 to 12V on the wire, a digital 0 bit as a voltage of negative −5 to −12V. The operation range of the bit rate is very large between 300 bit/s and 115,200 bit/s. The transmission rate must be precisely defined for both stations.

The application of RS-232 drivers and receivers in inductive sensors is advantageous due to the convenient parameters: output impedance of 50Ω, input impedance in a range of 3 to 7 kΩ [2], low current consumption, short-circuit protection of the outputs, and more parts in a package. As an example, STMicroelectronics offers the device ST232, compatible with MAX232 (manufactured by Maxim), which contains two drivers and two receivers in the housing.

For substantially more details on RS-232, reference [4] is a good source.

### 14.4.2 RS-422 and RS-485

At the usual bite rate of 19,200 bit/s, the maximum distance allowed by RS-232 is about 50m [2], which in an industrial environment can be a severe limitation. Further, the maximum data transfer rate can be a limitation for fast data acquisitions.

Interfaces such as RS-422 and RS-485 were developed to overcome these limitations. The RS-485 standard derives from the RS-422 and is defined by the layer 1/OSI model [8]. They basically differ through the number of users: RS-422 can support only two stations, whereas RS-485 allows a larger number of users.

RS485 is a balanced to ground interface and has only four signals: TD (transmit data) and RD (receive data) as well as the control signals: RTS (ready to send) and CTS (clear to send). It provides a handshake transmission of binary data in half-duplex communication, since a station cannot simultaneously transmit and receive data streams.

An increase in transmission speed and maximum cable length is achieved by using voltage differentials on two signal wires A and B. For a logic 0, the voltage level on line A is greater than on the line B by 5V, and vice versa for logic1.

Unlike RS232, which can provide a connection only between two stations, the RS485 allows up to 32 line drivers and 32 line receivers on one set of signal lines. This is achieved by tristate logic of the driver output. This pin can be at logic 0, logic 1, or high impedance. The last state effectively disconnects the driver from the line and is set by an enable/disable signal on the driver chip. Which station is allowed to transmit at what time is not specified in the standard and must be defined by a higher layer protocol (e.g., PROFIBUS).

The RS-485 driver is specified for a load impedance ≤ 54Ω that corresponds to 31 ULs (normal unit load), and so it can drive up to 31 receivers.

RS-485 works with TTL voltage levels and has substantially better transmission features: a data rate up to 10 Mb/s and a maximum distance of 1,200m (at 93.75 kbit/s).

To give a chip example, STMicroelectronics offers the devices ST485, which is a transceiver that contains one driver interconnected with one receiver in a package.

## 14.5 Synchronous Serial Buses for Sensors

Many ADCs or systems on chip (SOC) use a serial interface to communicate with a microprocessor or to a host computer. A concrete example is shown in Section 7.4.3.1, where the systems on chip LDC1101 is connected to the host MCU by means of an SPI interface (Figure 7.23).

Components such as ADC and SOC are currently implemented in up-to-date inductive sensors; hence the presentation of their interfaces is included in this book.

Local serial interfaces have the advantage of providing a processor-independent interface that does not affect processor wait states, bus hold time, or clock rates. The primary disadvantage is speed, because the data must be transferred bit-by-bit.

The serial interfaces are developed for character-oriented communication between MCUs and integrated peripherals, which typically cover short distances, are usually synchronous (data on one line and clock on another line) and have a variable number of signals (lines):

- 1-wire (e.g., SensorPath™ Bus/Texas Instruments);
- 2-wire (Inter-Integrated Circuit I$^2$C/defined by Philips Semiconductor);

- 3-wire (MICROWIRE™ Serial Interface/defined by National Semiconductor);
- 4-wire (Serial Peripheral Interface SPI/defined by Motorola).

### 14.5.1 4-Wire Serial Peripheral Interface

Serial peripheral interface (SPI), which was developed in 1987 by the semiconductor manufacturer Motorola (today NXP Semiconductors), is a simple, low-cost serial interface based on the master-slave principle, which theoretically has four signals [6, 9]:

1. Serial clock (SCLK);
2. Master out, slave in (MOSI);
3. Master in, slave out (MISO);
4. Slave select (SS) (low-active).

The master can provide a full-duplex connection with one slave. In this case, the connections are pin-to-pin (SCLK to SCLK, MOSI to MOSI, etc.).

In addition, the master can control more slaves. In this case, multiple slaves are connected in parallel to the master pins: SCLK, MOSI, and MISO. Individual SS wires are necessary to select the slave one we want to talk to.

The device can be used in four timing modes, in which the sampling time is dependent on the clock polarity and clock edge. For the common timing mode, the SCLK pulse is positive (same phase as the bit signal). Data is output on the rising SCLK edge and input data is latched on the falling edge (sample in reception) [9]. Clock frequencies up to the order of a few megahertz are possible.

The SPI data is device-specific. Some SPIs need a single byte; others need multiple bytes. Generally, reads follow writes. A dummy write is necessary to initiate a read (to shift the slave register data to the master).

The SPI interface is widely used in up-to-date complex devices as microcontrollers, SOCs, and the majority of ADCs (e.g., 8-bit AD7823/Analog Devices, 10-bit MAX1242/Maxim). The SPI macro is also available and can be integrated in new ASICs.

### 14.5.2 2-Wire Interintegrated Interface

The interintegrated ($I^2C$) interface was developed in 1982 by Philips Semiconductors (since 2006 NXP Semiconductors) and is today in the portfolio of theoretically all semiconductor manufacturers. The granted patent expired in 2006, so that $I^2C$ is licence-free. The bus is suitable for applications having components at different places and connected via a cable, rather than applications with devices at the same place.

It is a very simple, low-cost serial, master-slave interface that uses only two pins (wires) [10]:

1. Serial Clock (SCL), which is generated by the master to clock data into and out of the slave devices;

2. Serial Data (SDA), which is a bidirectional line that serially transmits all data into and out of the peripheral. The SDA pin is an open collector, so several users can share the same two-wire bus.

When sending data, SDA signal is allowed to change only while SCL is in the low state.

Transitions of SDA while SCL is high are interpreted by all slaves as START or STOP/END conditions (start for SDA going low and stop for SDA going high).

Every bus user has a unique address and the system can be implemented as multimaster with collision avoidance (see Section 14.1.2). Hence, the typical transfer data consists of (Table 14.3)

- Start bit.
- 7-bit address field to select which device will be accessed.
- Read/write bit (R/W) (1 for read, 0 for write). This bit refers to the addressed device and specifies if this has to be read or written. The selected user will drive the SDA line low to indicate that it has received the address and R/W information. After the R/W-bit, the devices program their SDA pins to be input or output, correspondingly.
- Acknowledge bit.
- 8-bit data field. For a write operation, the master clocks out 8 data bits, and for a read operation, the master treats the SDA pin as an input and clocks in 8 bits.
- Stop bit.

One drawback of I$^{2}$C is the speed. Initially developed for 400 kbit/s, the interface was continually improved. Today a high-speed version works up to 3.4 Mbit/s. There are no demands for an exactly defined transmission rate. High speed and fast mode support a 10-bit address field, and so they can address up to 1,024 users, but they need active pullups to limit the cable capacitance. They are backward-compatible with older systems.

Like the SPI, the I$^{2}$C bus has a wide application field. A powerful device PCF8591 (supplier NXP Semiconductors) can be used either as an 8-bit ADC or DAC and has an internal I$^{2}$C interface. The interface can be addressed via bus telegram but also by means of three address inputs A2 to A0. This hardware possibility allows up to eight devices to share a single I$^{2}$C bus [11].

For microcontrollers without I$^{2}$C option, the remote 8-bit I/O expander PC8574 [10] is a good solution to bidirectionally convert the 8 bits of a parallel microcontroller port into I$^{2}$C serial data.

**Table 14.3** I$^{2}$C Bus Telegram (Read Command as R/W = 1)

| | *Address* | | | | *Subaddress* | | | *R/W* | | *Data* | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Start bit | 0 | 1 | 0 | 0 | x | x | x | 1 | ACK | x | x | x | x | x | x | x | x | Stop bit |

## 14.6 AS Interface

As mentioned in the introduction of this chapter, AS Interface is a sensor/actuator network (www.as-interface.net) that is a widely used communication medium for inductive sensors.

AS Interface is clearly positioned on the bottom level: sensor and actuator level of the five-level CIM stack (Section 14.2) and below the middle production and process (fieldbus) level. Thus, it is not a competitor to a fieldbus (PROFIBUS, InterBus, DeviceNet, etc.), but rather an essential and economic supplement.

On the bottom level, mainly binary actuators and sensors (more than 80%) are networked. Many of the connected devices either supply or require binary signals.

As already shown before (Section 3.4), the majority of industrial sensors are proximity and displacement sensors (P&D) and—more important—the majority inside this immense market are inductive P&Ds. In this world, data volumes to be transferred are small, but the users' number and the speed of transmission are very high.

This is also the area where AS Interface comes into its own. Their motto is "Linking simple binary sensors and actuators with a single line, able to transport data and power."

### 14.6.1 AS Interface at a Glance

The project Actuator-Sensor-Interface (ASI) was launched in 1991 by an international society comprising 11 well-known sensor and actuator manufacturers. The first commercial AS Interface devices were already available on the market in 1994. The system is exclusively a digital interface for sensors and actuators, binary as well as analog, and can be upwardly kinked with fieldbuses or process control systems [12].

The benefits, such as the easy elaboration of an application without the need for in-depth system knowledge, simple installation and putting into operation, easy maintenance, and low costs, are attainable due to the following attributes, which were specific goals already set at the beginning of the project:

a. The complete transfer of data but also of the supply energy for all sensors and the majority of actuators occurs along a cable without special requirements: 2-wire, unshielded, without terminating resistors at the ends.
b. The data transmission is very robust; hence there are no restrictions regarding the network arrangement even for environments with high EMC disturbances. The AS Interface can be implemented for any conventional network topologies (Section 14.1.1): linear, star, or tree topology.
c. The connection to the interface contains reduced-volume electronics (single- or double-chip architectures). This coupling slave electronics is placed in dedicated and standardized coupling modules or panels or is inserted in the sensor body (network sensors). Figure 14.7 illustrates four combinations of sensors or actuators connected by means of a coupling module or with integrated coupling electronics. Intelligent sensors with integrated AS Interface

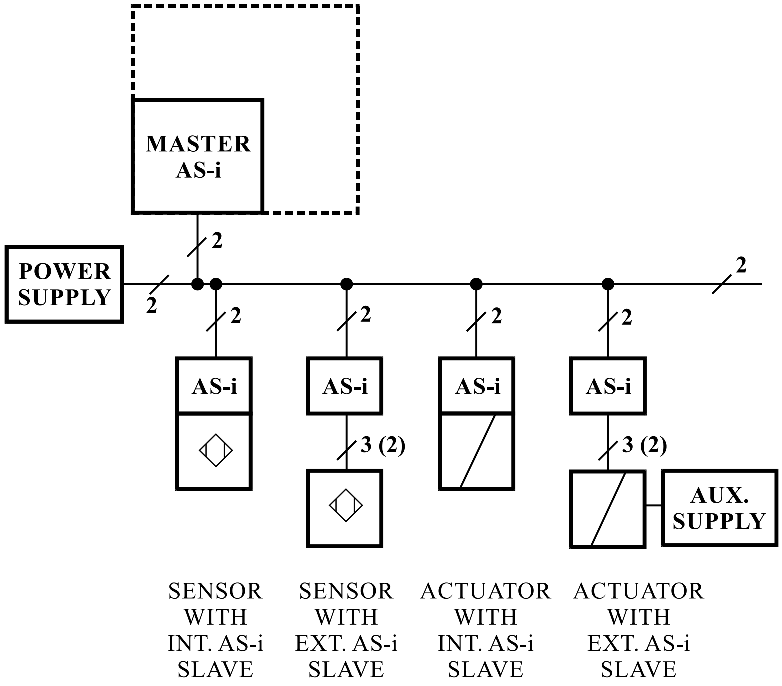

MASTER
AS-i
POWER
SUPPLY
2
2
2
2
2
2
2
AS-i
AS-i
AS-i
AS-i
3 (2)
3 (2)
AUX.
SUPPLY
SENSOR WITH INT. AS-i SLAVE
SENSOR WITH EXT. AS-i SLAVE
ACTUATOR WITH INT. AS-i SLAVE
ACTUATOR WITH EXT. AS-i SLAVE

**Figure 14.7** Block schematic of an AS Interface network containing the base components: master, power supply units, and sensors/actuators with external or internal slaves.

chips each have their own slave address and are identified as normal slaves by the master.

d. The system works with master-slave access and supports up to 31 slaves; the master can be implemented in computers or process control systems and can be connected to a RS232, RS422 interface or to a fieldbus.

A standard AS Interface system requires a supply unit (32 VDC/2A, typically), which is able to supply 124 sensors connected to 31 coupling modules (each module can drive up to four sensors). If actuators with high current consumption are also involved, they should be supplied from auxiliary supply units. Considering a cable with a length of 100m and with a wire cross section of 1.5 mm$^2$ (as recommended), the cable voltage drop should remain below 5V. Thus, the resulting minimal supply voltage of 27V is sufficient to supply the slave electronics and the sensors.

Up to 62 A/B slaves can be connected to an expanded AS Interface system, each with a maximum of four inputs and three outputs (i.e., up to 248 inputs and 186 outputs).

The modular concept and the specific cable are well designed. The yellow cable has become the AS Interface trademark. The cable has a quasi-trapezoidal cross section and therefore is resistant to polarity reversal. It can be connected to the coupling module at any place easily and safely. The module box is composed of two parts: (a) a mounting plate with a cable channel (negative shape of the cable) that forms the lower part, and (b) the module itself forming the upper part. The upper part contains the AS Interface electronics, as well as the connectors for sensors and actuators (usually M12 sockets) [12].

Figure 14.8 is a 3-D schematic representation of an AS Interface network having various components. It shows the cable-carrying mounting plate (a) and details the contact making by insulation piercing: the cable is put into the channel and if the upper part is applied there, its contact needles pierce through the cable insulation and make high safety contacts with the copper conductors of the cable. The solution is safe and versatile, since connections along a cable line can be shifted and replaced easily.

Additional auxiliary power is often required for actuators. A black cable with similar features was developed to enable the same installation technique in a second specific channel of the module.

### 14.6.2 AS Interface Slave Specification

The standard slave has two types of interconnection lines with sensors/actuators:

1. Data lines D0 to D3, which can be independently configured by master (input, output, or bidirectional) and can be interrogated.
2. Parameter setting lines P0 to P3 for immediate parametrization in the process (switching of operation modes, of value ranges, etc.). They are only slave outputs, can be independently controlled by the master, and are not considered to be data.

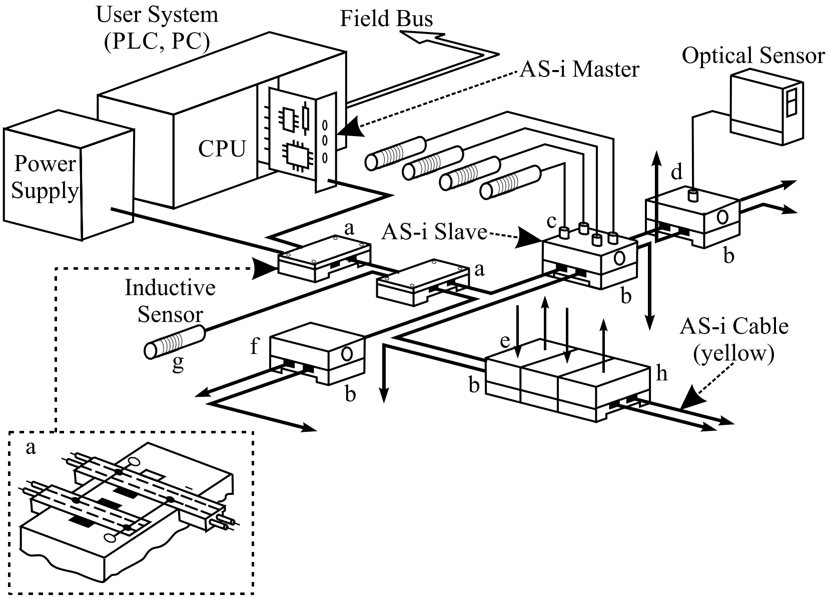


**Figure 14.8** A 3-D schematic representation of an AS Interface network. The magnified sketch (a) details the technique to make contact by insulation piercing [12].

Every slave is defined by means of the following three attributes:

1. Input/output Configuration, which fixes the type of data lines. The configuration is stored into an EEPROM memory of the slave. It is not available to the master; the only way to configure this 4-bit word is to program it by means of a programmer before the slave is connected to the bus. The 4 bits provide 16 binary combinations, which are insufficient to cover the spectrum defined by four data lines, where each line could have three valences: input, output, or tristate. The system specification includes the most used 16 combinations (e.g., 0000 describes a slave with four inputs, 0011 sets two inputs (D0, D1) and two outputs (D2, D3), 1000 defines a slave with four outputs, 1111 specifies four tristate lines, etc.) [12].
2. Identification code, which is a second 4-bit word, is also placed in slave memory and can be changed only by programming it at the slave manufacturer. The binary combinations of these two words make it possible to define 256 types of slaves, known as AS Interface slave profiles. To facilitate interchangeability and interoperability, these profiles are strongly standardized and controlled in the AS Interface community with respect to type and destination in the process of the data and parameter setting lines.
3. Slave address. AS Interface uses a direct addressing. Due to the reserved 5 bits for the address, the maximum number of addressable slaves is 31. Their addresses values are 1 ... 31 (or 1A ... 31A, 1B ... 31B for the extended specification) and can be programmed offline with an addressing unit before the slave is inserted in the network or changed online by the master in service. The address $0_{10}$ is reserved as an interchange buffer. The current address of a slave can be changed in two cycles:

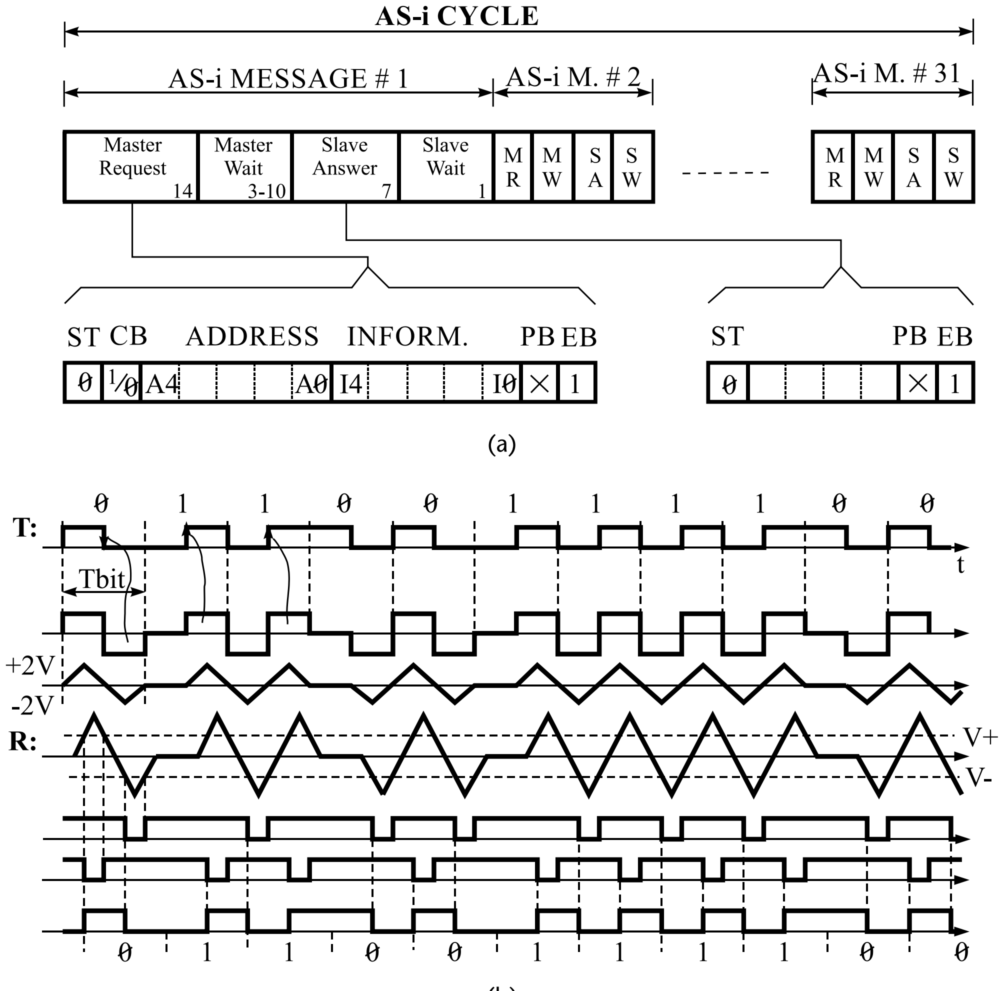
- First, the master sends a command to this slave to clear its existing address. The address is cleared and the slave temporarily gets the address $0_{10}$.
- In the next cycle, the master sends an address assigning a command specifying a new address. This instruction can be performed only by the slave with the temporarily assigned address $0_{10}$. The slave takes over the new address and releases the address buffer $0_{10}$. As long as the slave has the address $0_{10}$, it is unable to have other communications with the master.

### 14.6.3 ISO Reference Model of the AS Interface

To get compliant with the initial AS Interface target to replace traditional PtP connection by a data gathering bus with defined response time, the system was initially defined as a polling master-slave network (see Section 14.1.2).

The master provides the link with a higher-level control unit (PC, PLC, or gateway to a fieldbus) [13], and performs a cyclic and successive polling with the subordinated nodes (slaves), which are connected to the network at a given moment.

An AS Interface cycle contains a number of identical messages, equal to the number of slaves (Figure 14.9(a)). Because the cycle time duration is directly proportional


AS-i CYCLE
AS-i MESSAGE # 1
AS-i M. # 2
AS-i M. # 31
Master Request 14
Master Wait 3-10
Slave Answer 7
Slave Wait 1
M R
M W
S A
S W
ST CB ADDRESS INFORM. PB EB
0 1/0 A4 A0 I4 I0 1
ST PB EB
0 1
T:
Tbit
t
+2V
-2V
R:
V+
V-
0 1 1 0 0 1 1 1 1 0 0

(a)

(b)

**Figure 14.9** Overview of the ISO layers of the AS Interface: (a) ISO-layer 2 at maximal allocation, and (b) ISO layer 1.

to the slave's number, it is as short as possible and can be exactly evaluated (below 6 ms for maximal allocation).

An AS Interface message contains four binary words with fixed length:

1. Master request (14 bits). Beside the four single-bits: ST (start), CB (control), PB (parity control) and EB (end), this request consists of:
   - Address field (5 bits) for the slave address A4 to A0;
   - Information field (5 bits), whose data depends on the instruction type. It could be the words D0 to D3 or P0 to P3, a new address for the slave, or the code of a command call.
2. Master wait has a variable duration (3 to 10 bits/clocks) and is the waiting-state of the master. During this time, the slave tries to achieve the synchronism with the master and to confirm that by a reply after three or more

clocks. If after 10 clocks there is no answer, the master considers the slave to be defect, ignores it, and jumps up to the next slave.

3. Slave answer (7 bits). Besides the three single bits—ST (start), PB (parity control), and EB (end)—this answer has a 4-bit field with information to be communicated to the master: data or parameter values as well as the information required by the master regarding the slave; ID-code, content of its state register, etc.
4. Slave wait (1 bit) represents the rest time for the slave.

The valid ST and EB have the binary values 0 and 1, respectively.

The CB has two values: logic 0 for D0 to D3 or P0 to P3 or a slave address in the information field, and logic 1 for a command call.

The PB is 1 when the sum of the corresponding bits is 1 (even parity).

The system has a minimal set of nine commandos:

- Data exchange, which is the major commando and implies writing new data in the addressed slave and reading data from this slave;
- Parameter setting, where the master sets the bits P0 to P3 into the register of the addressed slave;
- Reset slave, which brings the slave into its default state;
- Clear address and assign address, which perform an address exchange (see Section 14.6.2);
- Four read commandos, which permit the master to read the I/O configuration, the ID-code, the status of the addressed slave or to read and erase the status register of this slave.

With regard to the IOS layer 1 description, the concept was defined to fulfill the following requirements: (1) modulated signals without DC component, (2) signal generation and transmission with minimal electronics (particularly on the slave level), and (3) narrow spectrum of the signal, because of the low cutting frequency of the chosen ordinary cable.

The clock frequency generation uses a quartz crystal with the frequency of 5.333 MHz. The standard clock frequency of 166.66 kHz results from dividing this frequency by 32 (easy binary division). Thus, the duration of a bit is $T_{bit} = 6\ \mu$s and the transmission rate is about 167 kbit/s, which is a reasonable value compared to the system simplicity. For the maximum configuration of 31 slaves, results a maximal AS Interface cycle duration of 5.95 ms.

To fulfill the three demands above, a new transmission procedure called alternating pulse modulation (APM) was elaborated. This is a combined modulation and consists of three steps (Figure 14.9(b)):

1. First, a similar code to Manchester II (Figure 14.4) is applied. A bit of 1 or 0 provokes a transition low-to-high or high-to-low in the middle of the interval $T_{bit}$.
2. The next step is an AFP modulation (Table 14.2), in which falling edges generate negative pulses (their duration is half of $T_{bit}$) and rising edges generate positive pulses (see the second time diagram in Figure 14.9(b)).

3. In the last step, the rectangular pulses are converted into a signal shape closed to the function $\sin^2$ (the result is simplified represented by triangles in the third time diagram). This conversion is beneficial because the signal has a narrow frequency spectrum and 95% of the signal energy is contained in the base spectrum, which is limited to 335 kHz (double-clock frequency)[12].

The signal processing at receiving is represented in the next diagrams. The received signal is compared with two threshold voltages: V+ and V–. The outputs of the comparators (see the fifth and sixth diagrams) are led to the inputs reset and set of a classical RS flip-flop. Its output reproduces the Manchester II code of the bits stream. A final sampling suffices to get the bits.

The theoretical analysis of the transmission safety concludes a low achieved Hamming distance [2] because the frames are too short. In order to increase the system reliability, specific actions are implemented in the hardware signal proceeding:

- Six successive samplings are made for every received Manchester II pulse;
- The receiver accepts only signals having the shape closed to $\sin^2$ function;
- Independent of the bit frame, the switching of the comparators alternate (observe the fifth and sixth diagrams);
- The change of frame lengths is an additional plausibility check;
- Finally, the parity check (see the two bits PB in Figure 14.9(a)) is involved.

As a result, applying these specific measures succeeds in achieving a high EMC immunity of the AS Interface components. Approval tests pass the specific EMC tests (see Section 2.2) with good results (severity levels 3 and higher). The system emission, including cable electromagnetic radiations, also fulfill the standards.

The AS Interface corresponds to the Euro Norm EN 50295 and to the international standard IEC 62026-2.

The system topology is very flexible and covers high distances for the area of sensors and actuators, generally placed closer to each other. For more than 100m, system repeaters (amplifiers) can be installed. The total length is limited to 300m. The cable can be lengthened with an extender.

Putting the system into operation is very easy. After applying the power supply, the master checks if the memorized network configuration matches the actual configuration. In case of coincidence, the master initializes the network and allows it to be used.

The identification and replacement of a defective slave is assisted by the master. It immediately recognizes a defective module, memorizes its attributes, isolates it, and signalizes the fault condition. The operator replaces the module by an operational slave with the address $0_{10}$. The master recognizes the new participant and tests its compatibility (I/O configuration and ID code). For positive results, the slave obtained the previous address and is integrated into the network.

The attitude of the semiconductor industry toward AS Interface was and is always attentive. On the market, there are more integrated circuits for AS Interface networks. The classical device is the AS2701 (manufacturer Austriamicrosystems) [14]. This device has:

- Two conditioned pins (LTGP and LTGN) for direct connection to the AS Interface cable;
- Eight I/O pins: D0 to D3 and P0 to P3;
- The connections for a 5.333-MHz XTAL crystal;
- A standard two-wire serial $I^2C$ interface (Section 14.5.2) for connection to an EEPROM to store the slave configuration. A small sized chip as M24C01 (manufacturer ST Microelectronics) with 1-kbit memory capacity in a miniaturized package is sufficient for the application.

Obviously, companies constantly improve their products. As an example, Austriamicrosystems now offers the AS2702 (upgrade of AS2701) with already integrated EEPROM.

Integrated Device Technology, Inc. (IDT) provides up-to-date AS Interface ICs as ASI4U, SAP51, and SAP5S with internal memory and LED drivers [15].

## 14.7 IO-Link an Up-to-Date Sensor Communication System

The IO-Link is the newest standard procedure to link sensors and actuators on the lowest communication level of the ISO/OSI stack. The idea of the merger all existing proprietary developments occurred in 2005, within the framework of the task group of the Profibus user organization [16].

### 14.7.1 IO-Link at a Glance

IO-Link device interface is an extension of the established three-wire 24V binary sensor interface (IEC 60947-5-2) (see Section 2.1.1) and not a fieldbus [16]. The connection topology between sensors/actuators and the inputs/outputs of coupling modules for different buses is a PtP type. The system is based on the master-slave principle in half-duplex mode.

The master may feature a number of ports for the connection of devices (slaves). Only one device may be connected to each port. The master is linked to an upper level control system: fieldbus (via a gateway) or backplane transmission channels within a PLC (Figure 14.10). At any time, only two users are linked in communication.

The general features and marketing benefits can be summarized as follows:

- Simple, robust and EMC immune interface;
- Reduced software demands, no need for device addressing;
- Backward compatibility with other existing systems;
- Interoperability (i.e., device independence of fieldbus manufacturer);
- Economical; IO-Link sensors are available at comparable prices as conventional sensors;
- Widespread acceptance in the world of sensors (developers, manufacturers, machine constructors, and users);
- International recognition, complying with IEC61131-9, 2013 standard: *Programmable controllers–Part 9: Single-Drop Digital Communication Interface For Small Sensors/Actuators* (SDCI).

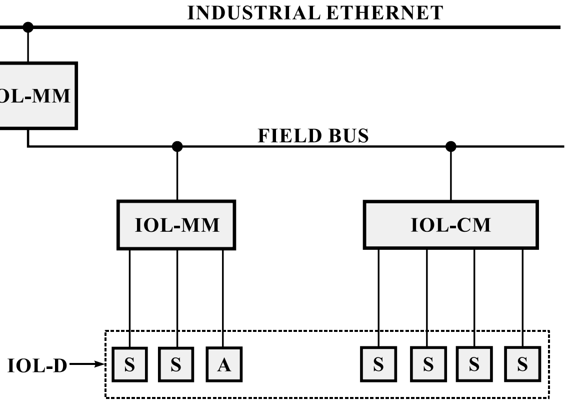


**Figure 14.10** Example of a system architecture with IO-Link: two types of IO-Link masters (master module IOL-MM and compact master IOL-CM) and several IO-Link devices IOL-D (sensors S and actuators A) [16].

The specific IO-Link functional features are as follows:

- The system performs digital transmission of sensor binary as well as analog information: online-process data, setting parameter, and events;
- For the connection of devices, the master provides two types of ports: type A for sensors and type B for actuators, and for three data transmission rates: COM1, 4.8 kbit/s; COM2, 38.4 kbit/s; and COM1, 230.4 kbit/s (Figure 14.11(a));
- Transmission method: pulse modulation at 24V.

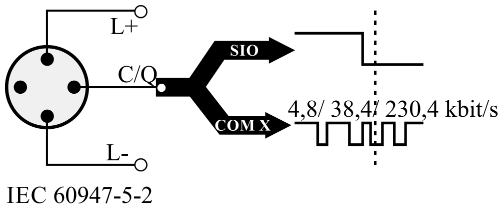
The technical parameters are very similar to the standard values for conventional sensors according to IEC 90647-5-2 (see Section 2.1.1):

- Supply voltage 24 V, device supply current ≤ 200 mA;
- Digital output current ≤ 200 mA;
- Transmission medium between two participants: unshielded three-wire cable, length ≤ 20m, wire cross section ≥ 0.34 mm$^2$;
- Cable loop parameters: $R < 6\Omega$, $C < 3$ nF at ≤ 1 MHz;
- Standard connectors (usually M12 size). The pins 1 and 3 are used for the supply voltage lines. The pin 2 (port type A) is optional and the pins 2 and 5 (port type B) are reserved for the auxiliary supply voltage of actuators (Figure 14.11(b)).
- The bivalent pin called C/Q (pin 4, black wired) is IO-Link-specific. It could sustain either a switching signal (standard input-output SIO) or a coded switching (communication line COM X).

### 14.7.2 ISO Reference Model of the IO-Link Interface

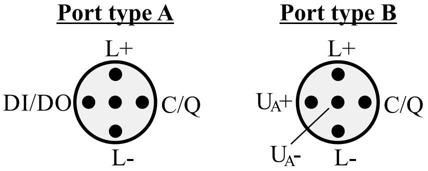
The serial network is developed for UART character-oriented communication between master and devices.

The IO-Link specification defines a frame as an informational unit comprising a master message and the subsequent device message (ISO layer 2), where the message



| Pin | Signal | Definition |
|---|---|---|
| 1 | L+ | 24 V |
| 2 | I/Q | Not connected, DI or DO |
| 3 | L- | 0 V |
| 4 | Q | “Switching signal“ DI,DO (SIO) |
|  | C | “Coded switching“ COM1, COM2, COM3 |

(a)



(b)

**Figure 14.11** (a) IO-Link standard connector with the general pin definition, and (b) port types and pining diagrams [18].

is the coherent, directed transfer of UART characters via IO-Link interface, from master to device or vice versa (ISO layer 1).

With respect to the ISO physical layer, for the bit-by-bit coding a NRZ (Figure 14.4) modulation is used.

A voltage of 0V between the C/Q and L– lines corresponds to the logic value of 1 and a voltage difference of +24V between the C/Q line and the L– line corresponds to the logic value of 0. The open-circuit level on the C/Q line is 0V with reference to L–.

The IO-Link-specific UART character is based on the standard Universal Asynchronous Receiver and Transmitter (UART) protocol described in Section 14.1.3.2. It has a length of 11 bits and is composed of

- ST (start bit) with the logic value of 0 (+24V between C/Q and V–).
- B0 to B7 (data bits). The second sequence bit B7 is MSB.
- PB (parity bit), which is the tenth sequence bit and is equal to 1 when the sum of the corresponding bits is 1 (even parity).
- SP (stop bit), the eleventh bit with the value 1.

Going to the ISO logical layer, the IO-Link frame consists of a master message and a device message. Each message has a fixed part (two characters for master and one character for the device) as well as a variable part, which depends on the data type:

1. On-request data, which is transmitted acyclically at the request of the application and includes parameter data or event data;
2. Process data, with high priority and in a fixed schedule. It includes measured or controlled variables. The transfer starts automatically following startup without the need for a request.

The best choice to explain the frame structure is the Frame Type 0 (see Table 14.4). It has the minimal length (four characters) and two different structures, depending on the master request:

1. For READ master request, the master message has two characters (CMD and CHK/TYPE) and is followed by the device reply message, which also has two characters (ON-REQ and CHK/STAT);
2. For WRITE master request, the master message is longer (three characters: CMD, CHK/TYPE and ON-REQ) and the device reply message is shorter (one character: CHK/STAT).

The character abbreviations are CMD = command; CHK/TYPE = check/frame type, ON-REQ = on-request data, and CHK/STAT = check/status.

To consolidate the understanding of the IO-Link frame and of the character structure, Figure 14.12 shows a general frame composed of more than four characters. The characters are shown in the middle of the figure. It is a READ master request and the device reply contains more process data characters.

To simplify the representation and to facilitate the understanding, the characters are represented by their significant bits (ST, PB, and SP were omitted).

Figure 14.12 also explains the bit fields of the fixed three characters: CMD, CHK/TYP, and CHK/STAT (see the explicative tables for these fields):

a. The bit R/W indicates the direction of the data via the data channel (i.e., read access from device to master or write from master to device).
b. The bits Data Channel of the CMD character show the data channel for access to the user data. The values of this parameter are listed in the related table. The value index service data unit (ISDU) is a service for the transfer of on-request data with acknowledgment. It is sent as a dedicated data quality on a dedicated data channel.
c. The next five bits contain the address of the user data on the specified data channel. If ISDU is selected via the data channel, these bits are used for flow control within the ISDU service. The address (i.e., the position of the user data within the service ISDU) is then only contained indirectly.
d. The first two bits of the CHK/TYPE-character specify the frame type, which is described below.
e. The second part of this character contains the bits called Checksum. This feature provides data integrity protection for data transmission in both directions. As mentioned above, each UART character is protected by an even parity bit PB. Besides this individual protection, all data bytes in a message are XOR (exclusive or) processed byte-by-byte. The resulting checksum bytes,

X7 to X0, are finally compressed in accordance with the following equations, from 8 to 6 checksum bits: CS5 to CS0 and entered into the characters CHK/TYP and CHK/STAT, respectively (Figure 14.12):

$$\begin{aligned}CS5 &= X7 \oplus X5 \otimes X3 \oplus X1\\ CS4 &= X6 \oplus X4 \otimes X2 \oplus X0\\ CS3 &= X7 \oplus X6\\ CS2 &= X5 \oplus X4\\ CS1 &= X3 \oplus X2\\ CS0 &= X1 \oplus X0\end{aligned} \tag{14.1}$$

f. The first flag called Event in the device reply CHK/STAT refers to an event, namely an on-request data quality. The transfer of events is initiated by setting this flag and takes place on the dedicated data channel. Events may be initiated in the application or in the communication protocol.
g. The second flag called Process Data indicates whether, given the conditions of an event, the device is able to make any valid process data available.

Depending on the data type and on the number of UART characters of variable part, the structures frame → message → character are classified according to [17] into three types with more subtypes:

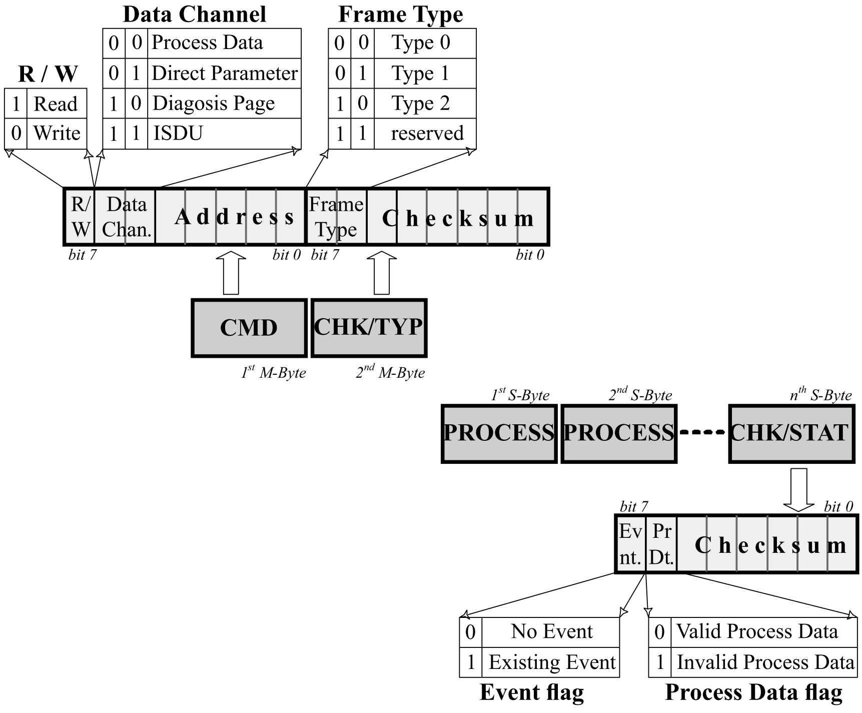

Data Channel
0 0 Process Data
0 1 Direct Parameter
1 0 Diagosis Page
1 1 ISDU
Frame Type
0 0 Type 0
0 1 Type 1
1 0 Type 2
1 1 reserved
R / W
1 Read
0 Write
R/ W
Data Chan.
A d d r e s s
Frame Type
C h e c k s u m
bit 7
bit 0
bit 7
bit 0
CMD
CHK/TYP
1st M-Byte
2nd M-Byte
1st S-Byte
2nd S-Byte
nth S-Byte
PROCESS
PROCESS
CHK/STAT
bit 7
bit 0
Ev nt.
Pr Dt.
C h e c k s u m
0 No Event
1 Existing Event
0 Valid Process Data
1 Invalid Process Data
Event flag
Process Data flag

**Figure 14.12** Block and detailed representation of a general IO-Link frame (see the middle range).

- Frame type 0, with a length of four characters (see the description above).
- Frames type 1, having two subtypes with a length of five characters (see Table 14.4).
- Frames type 2, with a length of four, five, or more characters. Table 14.4 contains only two examples. The types 2.3 to 2.6 are not represented but can be obtained from [17].
- The particular frames type 1.V and 2.V have a variable number of ON-REQ or ON-REQ/PROCESS characters, respectively. The maximal character number of these fields is 64 [16].

Note that although the newest version 1.1.2 of the IO-Link specification [18] is backward-compatible with the earlier version 1.0, it uses a changed terminology with respect to the initial version 1.0 [17] and the publication [16]. Thus, the frame is named M-sequence (to underline the master priority) and it consists of a master message and device message, and UART characters carry abbreviated names: MC (instead of CMD), CKT (≡CHK/TYPE), CKS (≡CHK/STAT), OD (≡ON-REQ Data), PD (≡PROCESS Data). The type classification remains the same.

**Table 14.4** IO-Link Frame Types [16]

| | | *Structures Frame → Message → UART Character* | | | | | |
|---|---|---|---|---|---|---|---|
| *Frame type* | | *UART Character #1* | *UART Character #2* | *UART Character #3* | *UART Character #4* | *UART Character #5* | *UART Character #6* |
| Type 0/R | —Master request | CMD | CHK/TYPE | | | | |
| | —Slave reply | | | ON-REQ | CHK/STAT | | |
| Type 0/W | —Master request | CMD | CHK/TYPE | ON-REQ | | | |
| | —Slave reply | | | | CHK/STAT | | |
| Type 1.1/R | —Master request | CMD | CHK/TYPE | | | | |
| | —Slave reply | | | PROCESS | PROCESS | CHK/STAT | |
| Type 1.1/W | —Master request | CMD | CHK/TYPE | PROCESS | PROCESS | | |
| | —Slave reply | | | | | CHK/STAT | |
| Type 1.2/R | —Master request | CMD | CHK/TYPE | | | | |
| | —Slave reply | | | ON-REQ | ON-REQ | CHK/STAT | |
| Type 1.2/W | —Master request | CMD | CHK/TYPE | ON-REQ | ON-REQ | | |
| | —Slave reply | | | | | CHK/STAT | |
| Type 2.1/R | —Master request | CMD | CHK/TYPE | | | | |
| | —Slave reply | | | ON-REQ | PROCESS0 | CHK/STAT | |
| Type 2.1/W | —Master request | CMD | CHK/TYPE | ON-REQ | | | |
| | —Slave reply | | | | PROCESS0 | CHK/STAT | |
| Type 2.2/R | —Master request | CMD | CHK/TYPE | | | | |
| | —Slave reply | | | ON-REQ | PROCESS0 | PROCESS1 | CHK/STAT |
| Type 2.2/W | —Master request | CMD | CHK/TYPE | ON-REQ | | | |
| | —Slave reply | | | | PROCESS0 | PROCESS1 | CHK/STAT |

### 14.7.3 IO-Link Communication Flow

Usually, an IO-Link device works in standard input-output mode SIO (see the simplified operating flow in Figures 14.11(a) and 14.13) but is ready for IO-Link communication.

The Wake Up event is characteristic for the IO-Link. Initiating this event, the master informs the device node, which is in either SIO mode or not active, that it will start an IO-Link communication. Therefore, the master drives the C/Q line to the opposite of its present state, and will either sink or source the wake-up current (<500 mA) for a wake-up duration (typically 80 $\mu$s). This controlled short circuit on the line is detected by the device node.

After this hardware action, the master sends a series of requests with the three transmission rates: COM3, COM2, and COM1. According to the specification, the device has to reply and to switch to receive mode within 500 $\mu$s after receiving the wake-up signal. If the device replies to one of these requests, the communication is established and the master reads further device parameters (direct parameter-page access). If the communication is not established, the master repeats the trial with the COM series and finally with low frequency (about 1 Hz).

The start-up phase is now established between master and device, during which a setup occurs (frame tape, specific times, etc.).

If the network complies with the version 1.0 of the IO-Link specification, the next phase is the cyclical mode, wherein the cyclical exchange of process data in a deterministic mode occurs for equally spaced time intervals. On-request data such as parameter values, event, and diagnosis data can be also transferred. The phase begins with a check of the device attributes (Vendor ID, Device ID).

If the network is in line with version 1.1 of the IO-Link specification, the check of the attributes occurs in a phase called preoperate. Using specific tools [18], the operator can set the device parameters. It can initiate transfers of service and request data but not process data yet. This setup occurs immediately and is shorter than setup actions during normal operation in cyclical mode.

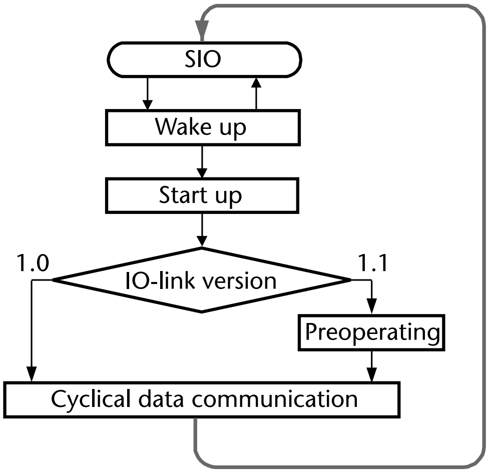


Figure 14.13 IO-Link communication flow [16].

The master can provoke a return to the SIO mode sending a fall-back command to the device. The device returns to SIO mode in no later than 500 ms, and the corresponding master port becomes a digital input.

### 14.7.4 IO-Link Hardware

For the three-wire interconnection of the master with a device, a specific transceiver TCVR circuitry between master CPU and the three-wire cable or between cable and device microcontroller is needed. This has to provide both operation modes: COMx as well as SIO (Figure 14.11(a)).

The IO-Link specification states a reference schematic, which is suitable for both sides (Figure 14.14). It contains two major stages:

1. A high-side output driver (called p-switching driver) connected as a transmitter to the bidirectional pin C/Q. Represented as a current source, it is characterized in active ON-state by a voltage drop *VR* and a standard current *IQ*. The source is controlled by an ON/OFF signal, which is delivered from the sensor electronics. An overcurrent event is indicated at the overload OVD output. As an option, there may be an additional low-side driver (n-switching driver) to allow a push-pull option.
2. An input comparator as a receiver, which converts the 24V signal on the C/Q line to standard logic level. It is a hysteresis comparator (see Section 9.6), which provides a switching signal H/L with high/low levels to the sensor electronics.

Concrete standard values of these parameters are available in the specification [18].

Similar to the AS Interface, the semiconductor industry has quickly responded to the need of integrated IO-Link transceivers, which act as complete physical layer for the communication.

To give only two examples, STMicroelectronics has designed and produces two competitive integrated circuits: L6360 (TCVR for masters [19]) and L6362A (TCVR for devices [20]). The device IC contains two MOS output transistors with separate outputs (which can be connected together or not), the input comparator and auxiliary stages as low-drop voltage regulator (3.3 or 5V), logic stage, overload detection, and standard protections (see Chapters 10 and 11). In addition, the master IC contains a powerful digital interface with the master CPU, an $I^2C$ bus (see Section 14.5.2.2) and two LED drivers.

The second example is given by TIOL111-x (Texas Instruments), an IO-Link device TCVR with the following core features [21]:

- High-side driver, low-side driver or IO-Link configurable output C/Q (Figure 14.15);
- COM1 to COM3 data rate support;
- Configurable limit of the current *IQ* = 50 mA to 250 mA (by $R_{SET}$);
- Wake-up current compliance;
- Low residual voltage *VR* < 1.75V at 250 mA;

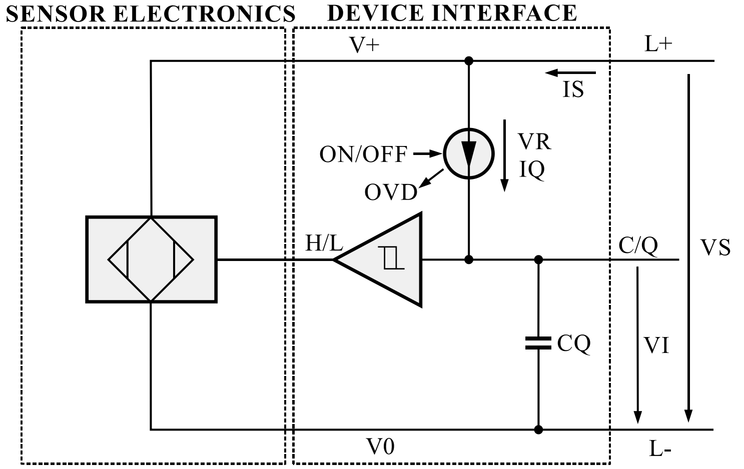

SENSOR ELECTRONICS
DEVICE INTERFACE
V+
L+
IS
ON/OFF
VR
IQ
OVD
H/L
C/Q
VS
CQ
VI
V0
L-

**Figure 14.14** Functional diagram of an IO-Link transceiver (signal symbols according to the IO-Link standard [18]).

- Low quiescent supply current < 1.5 mA;
- Large capacitive load driving capability, fast demagnetization of inductive loads (<1.5 H);
- Integrated reverse polarity protection and surge immunity up to 1.2 kV (500Ω) (see Chapter 11 and Section 2.2.4);
- Integrated voltage regulator of 3.3 or 5V for up to 20 mA current (version specific);
- Compact 10-pin VSON package (2.5 × 3 mm).

The block schematic in Figure 14.15 shows the standard IO-Link connections to the master and an easily understandable parallel interface to the sensor microcontroller:

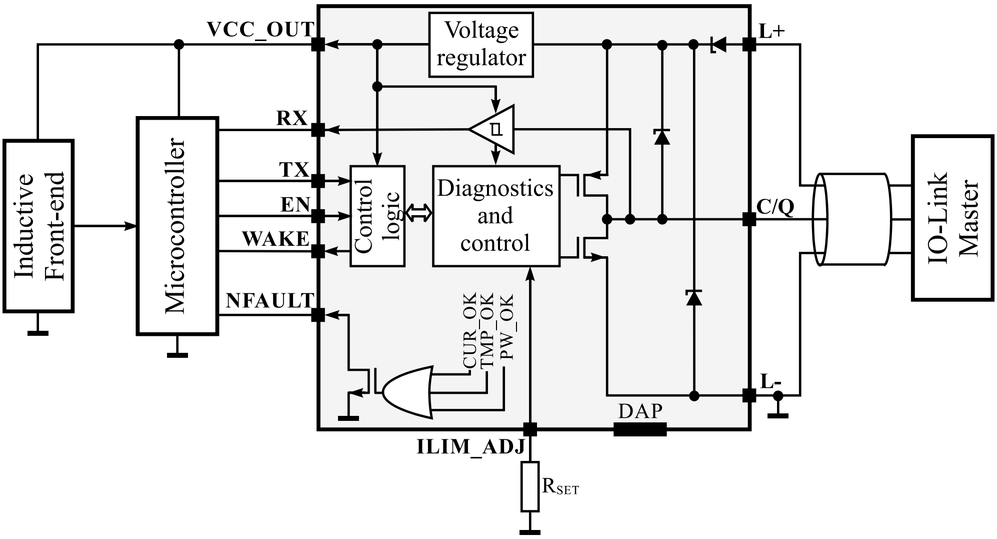

VCC_OUT
Voltage regulator
L+
RX
TX
EN
WAKE
NFAULT
Inductive Front-end
Microcontroller
Control logic
Diagnostics and control
CUR_OK
TMP_OK
PW_OK
C/Q
IO-Link Master
L-
DAP
ILIM_ADJ
RSET

**Figure 14.15** Block schematic of an inductive IO-Link sensor with integrated transceiver (application proposal for the integrated circuit TIOL111 [21]).

- VCC_OUT = regulated voltage for the sensor front-end.
- RX = received output data to the local microcontroller.
- TX = transmit data from the local microcontroller. The line is enabled when the signal EN is high. A logic high TX signal causes high-side driver switches on.
- EN = enable signal. Logic low of this signal sets the output C/Q at high-impedance.
- WAKE = open-drain wake-up signal to the local microcontroller.
- NFAULT = open-drain fault indicator to the local microcontroller. Its low level indicates (1) an overcurrent (signal CUR_OK), (2) an overtemperature (signal TMP_OK), or (3) an undervoltage < 6V supply (signal PW_OK) condition.

## 14.8 Conclusions

Trying a final benchmarking of most suitable data communication networks for inductive sensors for industrial applications, the conclusion is that the AS Interface covers intelligent applications having a medium complexity (hardware, software, installation) and offering medium transfer times. In constrast, IO-Link, which is not a bus system, is suitable for direct data transfers with low complexity but high communication speed. Both are adequate either for digital or for analog information.

An excellent system architecture that is actually proposed in IO-Link publications [16] is (1) to connect sensors via IO-Link to the ports of an IO-Link master and to link upward this master to an AS Interface bus, making it an AS-Interface slave (see the AS-i and IO-Link module AC5225, manufacturer ifm–electronics/Germany [22]).

### References

[1] Webster, J. G., *The Measurement, Instrumentation and Sensors Handbook*, Boca Raton, FL: CRC Press/IEEE Press, 1999.

[2] Schnell, G., and B. Wiedemann (eds.), *Bussysteme in der Automatisierungs- und Prozesstechnik: Grundlagen, Systeme und Anwendungen der industriellen Kommunikation, 8th ed*. Wiesbaden, Germany: Vieweg+Teubner Verlag/Springer Fachmedien GmbH, 2012.

[3] Tränkler, H.- R., and L.M Reindl, Sensortechnik–Handbuch für Praxis und Wissenschaft, Berlin, Germany: Spinger-Verlag, 2014.

[4] Nawrocki, W., *Measurement Systems and Sensors*, Second Edition, Norwood, MA: Artech House, 2016.

[5] Sze, S. M., *Semiconductor Sensors*, New York: John Wiley & Sons, 1994.

[6] Wilson J.,et al., *Test and Measurement–Know It All*, Burlington, MA: Elsevier/Newnes, 2017.

[7] http://www.analog.com/media/en/technical-documentation/data-sheets/aduc7019_20_21_22_24_25_26_27_28_29.pdf, data sheet, Precision Analog Microcontroller, 12-Bit Analog I/O, ARM7TDMI MCU, ADuC7019 ... 7029, 2005–2018.

[8] http://www.ti.com/lit/sg/slyt484a/slyt484a.pdf, RS-485 Reference Guide, Texas Instruments, 2014.

[9] http://www.ti.com/lit/ug/sprugp2a/sprugp2a.pdf, User Guide, Keystone Architecture Serial Peripheral Interface (SPI), Texas Instruments, 2012.

[10] http://www.ti.com/lit/ds/symlink/pcf8574.pdf, PCF8574 Remote 8-Bit I/O Expander for I2C Bus, Texas Instruments, 2015.

[11] https://www.nxp.com/docs/en/data-sheet/PCF8591.pdf, data sheet, PCF8591_8-bit A/D and D/A Converter, NXP Semiconductors, 2013.

[12] Kriesel, W., and O. W. Madelung (eds.) *AS-Interface Das Aktuator-Sensor-Interface für die Automation*, 2nd Edition, Leipzig, Germany: Fachbuchverlag, 1999.

[13] https://www2.ulpgc.es/descargadirecta.php, tutorial: AS-Interface, InterlinkBT, 2002.

[14] https://www.digchip.com/datasheets/parts/datasheet/060/AS2701.php, data sheet, AS2701 ASI Slave IC, AMS-Austria Microsystems, 2003.

[15] https://www.idt.com/document/dst/sap5-datasheet, Data sheet: SAP51 / SAP5S AS-Interface Compliant Universal AS-i IC, Spec. V3.0, 2016.

[16] Wienzek, P., and J. R. Uffelmann, *IO-Link–Intelligente Geräte brauchen einfache Schnittstellen*, Muenchen, Germany: Oldenbourg Industrieverlag, 2010.

[17] http://www.io-link.com/share/Downloads/Spec-Interface/IOL-Comm-Spec_ 10002 _V10 090118.pdf, IO-Link Communication Specification, version 1.0, Jan. 2009.

[18] http://www.io-link.com/share/Downloads/Spec-Interface/IOL-Interface-Spec_10002_ V112 Jul13.pdf, IO-Link Interface and System Specification V1.1.2, July 2013.

[19] www.st.com/resource/en/datasheet/l6360.pdf, L6360_IO-Link communication master transceiver IC–STMicroelectronics, May 2016.

[20] www.st.com/resource/en/datasheet/l6360.pdf, L6362A_IO-Link Communication Device Transceiver IC–STMicroelectronics, November 2017.

[21] www.ti.com/lit/ds/sllsev5a/sllsev5a.pdf, Date sheet: TIOL111 IO-Link Device Transceivers with Integrated Surge Protection—Texas Instruments, 2017.

[22] http://www.ifm.com/ifmca/web/news/as-i-module-with-2-io-link-ports.html, AC5225 Module/AS-I and IO-Link–A Good Connection, 2016.

# Key to the Symbols for Currents and Voltages Used Throughout the Book

To facilitate the study, the variety of these quantities (DC or AC, average or effective, etc.) are coded in this book using a notation with two alphanumerical digits, a letter and an index, and according to the following table.

As a general rule, capitals distinguish quantities invariable in time while small letters indicate time-variable currents or voltages.

Naming Conventions of Currents and Voltages

| | *Index* | |
|---|---|---|
| *Letter* | *Capitals* | *Small Letters* |
| *Capitals* | $I_C$, $V_{CC}$, $V_{BE}$ , and others<br>DC quantities and steady-state values | $I_c$, $V_{lc}$, and others<br>Effective and average values<br>$I_{cm}$, $V_{lcm}$, and others<br>Amplitudes |
| *Small Letters* | $i_C = I_C + i_c$, $v_{LC} = V_{LC} + v_{lc}$<br>Instantaneous values<br>$i_{CM}$, $v_{LCM}$<br>Maximum values, peak values<br>$i_{Cm}$, $v_{LCm}$<br>Minimum values, valley values | $i_c$, $v_{lc}$<br>Momentary value of the sinusoidal component (e.g., $i_c = I_{cm} \cdot \sin\omega t$) |

## Key to the Symbols for Electronic Components

To distinguish between practical (real) circuit components (resistor, capacitor, inductor, transistor, etc.) and their equivalent physical characteristics, the corresponding symbols are featured as following:

- Normal letters for practical part (for example, resistor $\mathrm{R}_1$, coupling capacitor $\mathrm{C}_\mathrm{C}$, inductor $\mathrm{L}_2$, transistor $\mathrm{T}_{11}$, etc.)
- Symbols in italic letters for physical magnitudes (for example, input resistance $R_i$, coil inductance $L$, base-emitter capacitance $C_{BE}$, etc.).

# Acronyms and Abbreviations

| | |
|---|---|
| 2D/3D | two-dimensional/three-dimensional |
| ABM | Asynchronous balanced mode |
| AcC | Acceptance criterion (EMC tests) |
| AC | Alternating current |
| A/D | Analog-to-digital |
| ADC | Analog-to-digital convertor |
| APLL | Analogue phase-locked loop circuit |
| APM | Alternating pulse modulation (AS interface specific) |
| APS | Analog proximity sensor |
| ARM | Asynchronous answer mode |
| AS-Interface | Actuator/sensor interface |
| ASIC | Application-specific integrated circuit |
| BCD | Binary-coded-decimal |
| BEM | Boundary element method |
| BJT | Bipolar junction transistor |
| BSL | Best straight line |
| BST-SG | Burst signal generator |
| CAD | Computer-aided design |
| CAE | Computer-aided engineering |
| CAGR | Compound annual growth rate |
| CAN | Controller area network |
| CCC | Capacitive coupling clamp |
| CDD | Coupling/decoupling device |
| CE mark | Communauté Europèene mark |
| CIM | Computer integrated manufacturing |
| CMCD | Common mode conducted disturbances |
| CMOS | Complementary metal-oxide-semiconductor |
| CMRR | Common-mode rejection ratio |
| CPU | Central processing unit |
| CRC | Cyclic redundancy check |

| | |
|---|---|
| CRV | Contact resistance variation |
| CSMA | Carrier sense multiple access |
| CSMA/CD | CSMA with collision detection |
| CSMA/AV | CSMA with collision avoidance |
| D/A | Digital-to-analog |
| DAC | Digital-to-analog convertor |
| DCE | Data communication equipment |
| DCN | Directional coupling network |
| DCS | Digital controlled current source |
| DMS | Distributed multidrop systems |
| DPLL | Digital phase-locked loop circuit |
| DSC_EL | Discharging electrode |
| DSP | Digital signal processing (processor) |
| DTE | Data terminal equipment |
| DUT | Device under test |
| EE | Evaluation electronics |
| EEPROM | Electrically erasable programmable memory |
| EFT | Electrical fast transients |
| EMC | Electromagnetic compatibility |
| EMI | Electromagnetic interference |
| ESD | Electrostatic discharge |
| ESD-SG | ESD-Signal generator |
| ESW | External switch (normal or bounce-free) |
| EVD | Evaluation device |
| EVU | Evaluation unit |
| FATM | File transfer, access, and management |
| FCS | Frame check sequence |
| FE | Finite element |
| FEA | Finite element analysis |
| FEM | Finite element method |
| FET | Field effect transistor |
| FIT | Failures in time |
| FM | Frequency modulation |
| FOP | False output pulse |
| FPGA | Field-programmable gate arrays |
| FR | Failure rate |
| FRAM | Ferroelectric random access memory |
| FRO | Full-range output |

| | |
|---|---|
| FSR | Full-scale range |
| GND | Ground |
| GTEM | Gigahertz transverse electromagnetic |
| GUI | Graphical user interface |
| H | Hysteresis, differential travel |
| HALT | Highly accelerated life test |
| HART | Highway addressable remote transducer |
| HF | High frequency |
| HLDL | High level datalink control |
| HSD | High-side driver |
| IC | Integrated circuit |
| IDS | Inductive displacement sensor |
| IEEE | Institute of Electrical and Electronics Engineers |
| IEC/CEI | International Electrotechnical Commission/Commission Electrotechnique Internationale |
| I/O | Input/output |
| IO-Link | Input/output link |
| IOL-D | IO-link device |
| IOL-MM | IO-link master module |
| IOL-CM | IO-link compact master |
| IP | International protection codes |
| IPS | Inductive proximity sensor |
| IPS_UT | Inductive proximity sensor under test |
| IPSW | Inductive proximity switch |
| IS | Inductive sensor |
| ISE | Inductive sensing element |
| IS_CML | Inductive sensors with closed magnetic loop |
| IS_EC | Inductive sensor based on eddy-currents |
| IS_OML | Inductive sensors with open magnetic loop |
| ISDU | Index service data unit (IO-link specific) |
| ISF | Intrinsic safety |
| ISO | International Standards Organization |
| JFET | Junction field effect transistor |
| KCL | Kirchhoff's current law |
| KVL | Kirchhoff's voltage law |
| LDO | Low-dropout voltage regulators |
| LF | Low frequency |
| Li-Qxx | Trademark Turck Company: Inductive linear position sensor |

| | |
|---|---|
| LPF | Low-pass filter |
| LR | Linearity range |
| LSB | Least-significant bit |
| LSD | Low-side driver |
| LUT | Look-up table |
| LVDT | Linear variable differential transformer |
| LwD | Lower distance |
| LwL | Lower limit |
| MCU | Microcontroller central unit |
| MFI | Magnetic field immunity |
| MIDS | Magneto-inductive displacement sensor |
| MIL Standards | Military standards |
| MMCP | Mixed-mode communication protocols |
| MOS | Metal-oxide-semiconductor |
| MS | Market share |
| MSB | Most-significant bit |
| MTBF | Mean time between failures |
| MTP | Multi-time programmable/programming |
| MTTF | Mean time to failure |
| MTTR | Mean time to repair |
| MW | Metal wall |
| NC | Normally closed |
| NCAP | Network capable application processor |
| NEC | National Electrical Code |
| NO | Normally open |
| NRM | Normal answer mode |
| NTC | Negative temperature coefficient (thermistor) |
| OSC | Oscilloscope |
| OTP | One-time programmable/programming |
| OUT | Output |
| OWP | Open wire protection |
| P&D | Proximity and displacement |
| PAP | Power amplifier |
| PCB | Printed circuit board |
| PDU | Protocol data unit (IO-link specific) |
| PLC | Programmable-logic controller (control unit) |
| PLCD | Permanent magnetic linear contactless displacement sensor |
| PLL | Phase-locked loop circuit |

| | |
|---|---|
| PMU | Power measuring unit |
| POR | Power-on reset |
| PP | Push-pull |
| PRC | Parallel resonant circuit |
| PS | Proximity sensor |
| PSoC | Programmable system on chip |
| PSRR | Power supply rejection ratio |
| PSW | Proximity switch |
| PSU | Power supply unit |
| PTAT | Proportional to the absolute temperature (sensor) |
| PTC | Positive temperature coefficient (thermistor) |
| PWM | Pulse-width modulation |
| R | Repeat accuracy |
| RAM | Random access memory |
| RDT | Resistance temperature detector |
| RFI | Radiated electromagnetic field interference |
| ROM | Read only memory |
| RPP | Reverse polarity protection |
| SAG-PAD | Sagentia-pad |
| SAR | Successive approximation register |
| SC | Scope/Oscilloscope |
| SCSW | Short-circuit switch |
| SCT | Schmitt trigger |
| SCP | Short-circuit protection |
| SDCI | Single-drop communication interface |
| SDS | Software-defined sensor |
| SE | Sensing element/sensor element |
| SGT | Signal generator |
| S/H | Sample-and-hold circuit |
| SI | International System of Units |
| SIO | Standard input-output mode (IO-link specific) |
| SL | Severity level (EMC tests) |
| SMD | Surface mount device |
| SMT | Surface mount technology |
| SPS | Sensor power supply |
| SPI | Serial peripheral interface |
| SR | Sensing range |
| SRAM | Static random access memory |

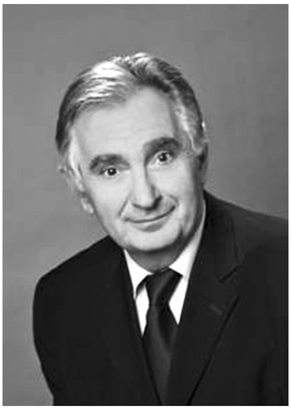
| | |
|---|---|
| SRC | Series resonant circuit |
| SRD | Shunt regulator device |
| SRG-SG | Surge signal generator |
| SSI | Serial synchronous interface |
| SSOP | Shrink small outline package (for ICs) |
| STG | Standard target |
| STIM | Smart transducer interface |
| SW | Switch, switching element |
| TCR | Temperature coefficient of resistance |
| TCVR | Transceiver |
| TDO | Tunel diode oscillator |
| TEDS | Transducer electronic data sheet |
| TG | Target |
| TII | Transducer independent interface |
| ToffD | Turn-off delay |
| TonD | Turn-on delay |
| TVS | Transient voltage suppressor |
| UART | Universal asynchronous receiver and transmitter |
| UL | Underwriter's Laboratories |
| UpD | Upper distance |
| UpL | Upper limit |
| USB | Universal Serial Bus |
| USD | U.S. dollar |
| VCO | Voltage-controlled oscillator |
| VDI | Voltage dips and interruptions |

# Nomenclature of Electromagnetic Quantities

| *Symbol* | *Quantity* | *SI Unit Name* | *Unit Symbol* |
|---|---|---|---|
| $\mathbf{A}$ | Magnetic vector potential | Tesla · m (weber/m) | T · m (Wb/m) |
| $\mathbf{B}$ | Magnetic flux density/ induction (vector) | Tesla (weber/$m^2$) | T (Wb/$m^2$) |
| $B$ | Susceptance | Siemens | S |
| $\mathbf{D}$ | Electric displacement (vector) | Coulomb/$m^2$ | C/$m^2$ |
| $\delta$ | Skin depth | Meter | m |
| $e$ | Euler's number (constant) | 2.71828 … | |
| $\mathbf{E}$ | Electric field intensity (vector) | Volt/m | V/m |
| $E$ | Electromotive force (emf) | Volt | V |
| $\Phi$ | Magnetic flux | Weber (Wb) | Wb (Vs) |
| $\Phi_E$ | Electric scalar potential | Volt | V |
| $\varepsilon_0$ | Absolute permittivity of free space | $8.854 \cdot 10^{-12}$ F/m | (farad/m) |
| $\varepsilon_r$ | Medium's relative permittivity | Nondimensional | |
| $\varepsilon$ | Permittivity ($\varepsilon = \varepsilon_0 \cdot \varepsilon_r$) | Farad/m | F/m |
| $G$ | Conductance | Siemens | S |
| $\mathbf{H}$ | Magnetic field intensity (vector) | Ampere/m | A/m |
| $\mathbf{J}$ | Electric current density (vector) | Ampere/$m^2$ | A/$m^2$ |
| $L$ | Inductance, inductivity | Henry | H |
| $\mu_0$ | Absolute permeability of free space | $4\pi \cdot 10^{-7}$ H/m | (henry/m) |
| $\mu_r$ | Medium's relative permeability | Nondimensional | |

| *Symbol* | *Quantity* | *SI Unit Name* | *Unit Symbol* |
|---|---|---|---|
| $\mu$ | Permeability ($\mu = \mu_0 \cdot \mu_r$) | Henry/m | H/m |
| $\mathbf{M}$ | Magnetization (vector) | Ampere/m | A/m |
| $\mathbf{P}$ | Polarization (vector) | Coulomb/m$^2$ | C/m$^2$ |
| $\pi$ | Pi number (constant) | 3.14159 ... | |
| $\rho_v$ | Volume charge density (scalar) | Coulomb/m$^3$ | C/m$^3$ |
| $\rho$ | Medium's resistivity | Ohm meter | Ωm |
| $R$ | Resistance | Ohm | Ω |
| $R_m$ | Magnetic reluctance | 1/henry | 1/H |
| $\sigma$ | Medium's conductivity ($\sigma = 1/\rho$) | Siemens/m | S/m |
| $\mathbf{V}$ | Carrier velocity (vector) | Ampere · m/coulomb | A · m/C |
| $V_m$ | Magnetomotive force (mmf) | Ampere · turns | |
| $\chi_\mathrm{e}$ | Electric susceptibility | Nondimensional | |
| $\chi_\mathrm{m}$ | Magnetic susceptibility | Nondimensional | |
| $W$ | Energy | Joule (watt · second) | J (W · s) |
| $X$ | Reactance | Ohm | Ω |
| $\underline{Y}$ | Admittance (complex quantity) ($\underline{Y} = G + \mathrm{j}B$) | Siemens | S |
| $\underline{Z}$ | Impedance (complex quantity) ($\underline{Z} = R + \mathrm{j}X$) | Ohm | Ω |

# About the Author



Dr. Sorin Fericean is a professional expert on the subject of design and manufacturing of electronic sensors for industrial applications. Up until 2015, he had been engaged in business for more than twenty-five years with one of the ten global players on the field of industrial sensors, that is, Balluff GmbH Company, Germany. He started here as a sensor design engineer and completed inductive sensor development projects. Later, Dr. Fericean was promoted to become head of a development department, responsible for company ASIC design projects, in cooperation with several IC design houses and manufacturers in the U.S. and Europe.

In 2006, Dr. Fericean assumed responsibility for the then newly established research and development department at Balluff named the Company Innovation Management Department, with two core competences: designing and prototyping of sensors based on novel functional principles, as well as innovation and intellectual propriety management.

Since his retirement in 2015, Dr. Fericean has been working as a freelancer in his consulting office FerSensC/Germany, carrying out sensor design projects.

His research fields are sensor engineering in general and inductive sensors in particular, ASIC designing, testing and release for serial production, as well as microwave-based sensors.

Dr. Fericean is an IEEE senior member and has actively and repeatedly participated in IEEE conferences on sensors as well as in other scientific events in the U.S. and Europe.

He has coauthored two volumes on sensors published in Germany, and has appeared as author or coauthor of 40 professional articles (issued in IEEE, Elektronik, Technisches Messen, etc.). His inventiveness shows in 34 patent families (70 total patent drafts), the vast majority of which have already been granted and implemented on a large scale.

# Index

Absolute voltage value circuit, 332
AC/DC signal convertors, 329
  Full-wave (precision) rectifier, 329, 331–333
  Half-wave (precision) rectifier, 329, 330–331
Acceptance criteria, 37
Accuracy, 9
Active face, 3
Active temperature compensation, 451
Active temperature-dependent circuits
  Bandgap cell, 446
  Bandgap references, 444–446
  Bipolar transistor sensors, 439–442
  Proportional to the absolute temperature (PTAT), 442–444
Actuation, 4
  *See also* Contactless Actuation
Adjustment resistor, 266
Admittance, 95, 98, 239
Ampere's law, 135
Amplifier
  Adder-substractor, 317
  Exponential, 317
  Instrumentation, 319
  Inverting, 261
  Logarithmic, 317
  Noninverting, 261
  Summing, 317
Amplifier gain, 252
Analog front-end, 209
Analog hardware-based compensation 452–456
Analog-to-digital convertors (ADC)
  Dual-slope ADC, 370–371
  Flash ADC, 368–370
  Sigma-delta ADC, 370–371
  Successive approximation ADC, 368–370
  Tracking ADC, 368–369
Analyzers
  Impedance, 109
  Vector network, 109
Angular frequency (pulstance), 243, 254
Aperture time, 342
Application-specific integrated circuit (ASIC) 231
Application-specific intelligent sensor (ASIS) 231
Approach
  Axial, 21
  Bilateral, 6
  Lateral, 21
  Unilateral, 6
Aquisition time, 343
Archimedean spiral, 146
Artificial Neural Network (ANN)
  Hidden, 352
  Neuron, 352
  Weight values, 353
AS Interface/Expanded AS Interface system 476, 482–489
AS Interface alternating pulse modulation (APM), 486–488
AS Interface cycle, 486
AS Interface ISO Model, 485
AS Interface message, 486
AS Interface slave profiles, 485
AS Interface repeaters, 488
AS Interface slave specification, 484
AS Interface trademark
  Yellow cable, 483
AS-i und IO-Link module AC5225, 498
Availabilty (AV), 48
Avarage voltage-value circuit, 333

Balloon models, 123
Bandwidth, 11, 108, 256, 415
Barkhausen criterion, 256
Base processing, 199
Basic amplifier, 251
BH curve, 165
Boundary conditions, 123
Boundary elements, 119
Brocaw call, 446
Bumping process, 204
Bumps
  Gold, 204
  Soldering, 204
Buried layer, 198

Calibration
  Parasitic, 134
  Process, 344
  Sensor, 188, 407
Capacitance
  Teach-in, 188
Capacitive-resistive behavior, 238
Carbonyl iron, 180
Central station (network), 467
Character-oriented communication, 479
Charge carriers
  n-type/negative charge/electrons, 194
  p-type/positive charge/holes, 194
Chip on board, 203
Chopping signal/equation, 337, 338
Clamping voltage, 209
Clean and low jitter signal, 225
Closed magnetic loop, 58
Closed-loop gain, 252
CMOS transmission gate, 342
Coding digital information
  Alternating flanks pulses (AFP), 472
  Bipolar, 472
  High-density bipolar (n=2) (HDB2), 472
  Manchester II (phase modulation), 472
  Nonreturn to zero (NRZ), 472
  Nonreturn to zero insert (NRZ I), 472
  Return to zero (RTZ), 472
Coercive field strength, 164, 165
Coercivity, 164
Coil
  Cylindrical, 133
  Flat planar, 133
  On silicon/on chip, 133
  On substrate/on FR4-PCB, 133
  Printed flat spiral, 143
  Wire-wound, 136
  With core, 133
  Without core, 133
Coil losses
  Additional losses, 141, 145–146
  Copper losses, 141–142
  Iron losses, 141, 143–145
Combustion triangle, 46
Communication
  Wired, 16
  Wireless, 16
Communication network topologies
  Circular network topology, 467
  Linear network topology, 466
  Point-to-Point (ptP), 466, 489
  Star network topology, 467
  Tree network topology, 466
Comparator
  Hysteresis, 311
  Threshold, 113
  Voltage, 353
  Zero-cross, 265
Compensation techniques
  Lag compensation, 326–327
  Lead compensation, 326–327
  Lead-leg compensation, 326, 328
  Modified lag compensation, 326–327
Complex plane representation, 100, 101
Computer integrated manufacturing (CIM)
  Sensor/actuator level, 474
  Production and process level, 474
Computer-aided engineering (CAE), 114
Computer-assisted analysis, 114
Condistdist®, 389
Conductance, 98, 239
Conductive
  Pads, 148
  Tracks, 148
Conductive paths (tracks) processing, 199
Conductivity, 164
Conformity (standard), 34
Connection, 4
Contact
  Electrical, 3

  Mechnical, 3
Contacts processing, 199
Convertor
  Electrodynamic, 52
  Electromagnetic, 52
  Electrostatic, 52
  Electrostriction, 52
  Piezoelectric, 52
Core
  Forming, 175
  Gaped, 169
  Inductance factor (AL-value), 170
  Milling/granulation, 175
  Permittivity, 171
  Resistivity (bulk), 170
  Sintering, 175
Core effective parameters
  Effective length, 139
  Effective surface, 139
  Effective volume, 139
Core factor, 139
Correction
  Zero/span, 344
Counter, 113
Coupling unit, 469
Critical dimension, 195
Curie temperature, 171, 173
Current
  Minimum operating, 27
  No-load supply, 27
  OFF-state, 27
  Rated operational, 27
  Source, 215
Current amplification, 315
Current consumption, 5
Current excitation model
  Solid, 123
  Stranded, 123
Current mirror, 271, 446
  Multioutput, 280
  Widlar, 276
  Wilson, 276
Current regulator diode (CRD), 208, 311
Current sink, 207
Current source
  Bidirectional, 285
Current-controlled voltage/current source, 315
Curve
  Characteristics, 6
Curve fitting, 347
Curve shaper, 345, 346
Cut-in diode voltage, 330

Darlington transistor, 208, 385
Data Error Handling/Hamming distance, 470, 488
Data frames rotation, 470
Data retention, 415, 417
Data transmission, 3
Deflection bridge, 65
Demagnetization, 165
Device
  Emission, 36
  Immission (susceptibility), 36
Device under test (DUT), 24
Die attached pad (metallic) (DAP), 226
Die/dice, 200
Differential travel, 24
Differentiator, 318, 319
Diffusion, 206
Digital potentiometers
  With volatile memory, 415
Digital Temperature Compensation, 456–463
  Actuating unit, 458, 459
  Communication and programming unit, 458
  Control unit and peripherals, 458
  Temperature management unit, 458, 459–461
Digital-to-analog convertor (DAC), 313, 365
  Binary-weighted DAC, 366
  R–2R ladder DAC, 367
Diode
  Array, 209
  Lambda, 292, 293–294
  Tunnel (Esaki), 292–293
Dirac delta function, 301
Distance range
  Lower distance, 33
  Upper distance, 33
Dopant
  n-type/donor impurity/electrons, 195
  p-type/acceptor impurity/holes, 195

Drift
  Long-term, 10
  Short-time, 10
  Span, 11
  Temperature, 11
  Thermal, 11
  Zero, 11
Droop rate, 343
Duty cycle, 225

Early voltage (VA)/effect, 274, 440
Eddy currents, 66
  Inductive sensors based on Eddy currents, 74
EEPROM, 415
Electromagnetic disturbances, 36
Electrical data domain (of sensor output), 373, 465
  Analog domain, 465
  Digital domain, 465
  Time domain, 465
Electromagnetic Compatibility (EMC), 30
Electromagnetic field
  Focusing/collimating, 137
  Orientation/guidance, 137
  Shielding, 137
Electromagnetic Interference (EMI), 35
Electromagnetic perturbations
  Unintentional perturbations, 36
  Volitional perturbations, 36
Element
  Sensing, 3
  Sensor, 3
EMC standards
  Basic, 36
  Generic, 36
  Product/product family, 36
Emitter follower, 215
Emitter processing, 199
End-to-end resistance, 412
Envelope detection, 113
Epitaxial (epi) growth, 198
Equation inversion, 347
Equivalent parallel resistance, 105
Error
  Absolute linearity, 9
  Linearity, 343
  Offset, 343
  Sensitivity, 343
  Total, 9
European hazardous zones
  Zone 0 to 2, 46
Evaluation electronics, 3
Evaluation electronics
  Discrete, 190
EX applications/devices, 46
Explosion proofing, 46

Face-down attachment, 204
Facedown die attachment, 157
Face-up attachment, 203
Faceup die attachment, 157
Failure rate (FR), 47
Failures in Time (FIT), 47
Feedback network, 252
Feedthrough, 342
Ferrite, 163
  Metallurgy, 163
  Sintered, 163
Fiel plots, 127
Field calculator, 128, 130
Field simulation
  Tools, 114
Fieldbus topologies
  Circular network topology, 470
  Linear network topology, 469
File transfer, access and management (FATM) 474
Final testing, 203
Finite elements
  Tetrahedrons, 119
  Triangles, 119
Finite-state machine, 189
Fourier series representation, 266, 271, 284, 300
Four-wire connection technique, 439
Free zone, 21
Frequency
  Corner, 254
  Cutoff, 166
  Divider, 312
  Gain-crossover, 325
  Of operating cycles, 28
  Phase-crossover, 325
  Self-resonance, 134
Frequency-response characteristic, 100

  Asymptotic Bode diagram (Bode plot), 320–323
Full-duplex fashion, 478
Full-range output (FRO), 6
Full-scale range (FSR), 6
Function generator, 305

Gain
  Amplifier, 254
Gain-bandwidth product, 266, 323
Gilbert cell multiplier, 341
Glob-top, 203
Graphical user interface (GUI), 120, 228–230
Green part, 175
Gyrator/active gyrator circuitry, 158

Half-duplex fashion, 479
Harmonic oscillator, 243
Highly accelerated life test (HALT)
  Test-to-fail, 49
High-side Driver (HSD), 13
Highway Addressable Remote Transducer Interface (HART), 476
Hold interval, 342
Holding capacitor, 342
Hyperbolic tangent function (tanh(x)), 297–298
HyperProx, 462
Hysteresis, 6
  Constant, 169
  Loss constant, 170
Hyteresis loop (magnetic), 164

IC
  Manufacturer/end user, 194
  Monolithic, 193
IC housing
  Ceramic, 200
  Plastic, 200
Ideal feedback amplifier, 251
IEC 60947-5-2, 20–32
IEC 60947-5-7, 32–35
IEC 61000-4-xx, 36–42
IEEE 1451.x, 476
Imaginary variable, 254
Immittance, 95, 98
Immunity
  Common mode conducted disturbances, 37, 41
  Electrical fast transients (EFT), 30, 37, 39
  Electromagnetic field (RFI), 30, 37, 38
  Electrostatic discharge (ESD), 30, 37, 38
  Impulse voltage/Lightning surges, 30, 37, 40, 209, 402–406
  Power frequency magnetic field, 37, 42
  Radio frequency interference (RFI), 37, 38
  Voltage dips and interruptions, 42
Impedance, 95
  Dip, 106
  Mutual, 97, 118, 153
  Peak, 106
  Self, 95, 118, 153
Inductive sensing element, 395
Inductive-resistive behavior, 238
Inductor
  Active, 158
  Linear variable, 71
  Perfect, 97
  Real, 97
Initial magnetization (virgin) curve, 165
Ink mark, inking, 200
Input overdrive, 384
Integral curves, 247
Integrated circuit
  Commercial, 2
  Manufacturer-proprietary, 2
Integrated trimming device, 416
Integration process
  Analog/digital mixed-signal process, 205
Integration process (planar)
  Back end, 194
  Front end, 194
Integration technology, 193
  Bipolar, 194
  Bipolar CMOS (BiCMOS), 206
  Complementary MOS (CMOS), 205, 206, 223
  High-power double-diffused (DMOS), 206
  Large-scale, 193
  Metal-oxide-semiconductor (MOS), 205
Integrator, 318, 319
Intelligent sensor/Smart sensor, 475–477

Interdependent system quantities, 247
Interface
  Serial peripheral (SPI), 16
  Serial synchronous (SSI), 16
International IP-Classification, 45
Intrinsic safety, 45
Inverse Laplace transform, 254, 301
IO Link, 476, 489–498
IO Link data transmission rates (COM1 to 3), 490
IO Link device, 489
IO Link ISO communication flow, 495
IO Link ISO frame types, 494
IO Link ISO hardware, 496
IO Link ISO Model, 490–494
IO Link master, 489
IO-Link, 420
Ion implantation, 206
Iron/iron (soft ) alloy, 137, 185
Iron alloy
  Nickel, 185
ISO/OSI Network model layers
  Application layer, 474
  Data link layer (logical layer), 474
  Network layer, 474
  Physical layer, 473
  Presentation layer, 474
  Session layer, 474
  Transport layer, 474
ISO/OSI Network reference model, 471, 473

Junction field effect transistor (JFET), 208

KCL/KVL, 277, 288
Kick-back/flyback pulse, 383
Kickback glitches, 219
Kinetik/potential energy, 242
Kitpart, 200

Laplace operator, 254
Large scale integration, 193
Level shifter, 228, 271
Life time, 4
Light emitting diode (LED), 208
Linear system with negative friction, 243
Linearisation
  By f-inverse, 351–352
  Logic hardware, 349
  Methods, 343
  Process, 344
  Signal, 343
Linearity
  Error, 7
Look-up table (LUT), 313, 347–349, 463
Loss factor, 168
  Coil, 79
  Relative, 169
Loss factor/loss angle of a coil, 102
Losses, Ohmic, 65
  Eddy (Foucault) current, 65, 169
  Magnetic aftereffect, 65
  Magnetic hysteresis, 65, 169
  Residual, 169
Lossless/lossy resonant circuit, 248
Low limit
  Live zero, 33
  True zero, 33
Low-side Driver (LSD), 13
Low-voltage drop regulator (LVDR), 227
Lumped inductances, 250

Magnetic constant of free space, 166
Magnetic losses, 61
Magnetic material
  Disaccomodation coefficient, 173
  Disaccomodation factor, 173
Magnetic vector potential, 117
Magnetic Yoke, 59
Magnetization, Saturation, 164
Majority carriers, 195
Manual mechnical potentiometers, 411–413
  Linear/logarithmic, 412
  Potentiometer/rheostat modes, 411
  Slide/rotary, 412
Margin, Gain/phase, 325
Materials
  Damping, 21
  Nondamping, 21
Matrix
  Inductance, 118
  Resistance, 118
Maxwell
  Equations, 114
  James Clerk, 114

MAXWELL electromagnetic field simulation (software), 59, 64, 69, 70, 74, 77, 127–130, 137, 145–146, 171, 176, 269
MC100EL1648, MC1648, 276
McLaurin series, 62
Mean time between failures (MTBF), 47
Mean time to failure (MTTF), 47
Mean time to repair (MTTR), 48
Measurand, 2
Mechnical linkage, 4
Mechnical potentiometer layers
  cermet conductive substrate, 412
  conductive plastic, 412
  wire wound, 412
Memory types
  One-time programable (OTP), 415, 418
  Multitime programable (MTP), 415
Metal deposition, 199
Metal sputtering, 199
Metal-Resistor Thermometers
  Nickel and Platinum, 435, 438–439
MiniDist®, 388
Mobile plunger
  Inductive sensors with, 66
Model mesh
  Mesh refining, 125, 126
Modulation
  Pulse-width (PWM), 16, 419
Moore’s law, 195
Multilayer perceptron (MLP), 352
Multivibrator
  Astable, 307, 311, 387
  RC-based (flip-flop), 306
Mu-metal (anneled), 137, 182

National Electrical Code (NEC), 46
Natural exponential function, 267
NEC hazardous locations
  Class I to III, 46
Negative feedback, 253
Negative temperature coefficient thermistor (NTC), 435–438
Neper frequency, 254
Network access procedures:, 467
  Master-slave procedure, 467
  Token procedure, 468
  Token-passing procedure, 468
Nonconductive substrats
  Epoxi resin FR4, 148
  Polymide for flex-PCBs, 150
  Prepregs, 149
  Rogers_RO4003, 150
Nonelectric measurands, 53
Nonlinearity, 7
  Endpoints, 8
  Independent, 7
  Zero-based, 8
Nonrecurrent expenses (NRE), 211
Nonvolatile, 5
Normally closed (NC), 13
Normally open (NO), 13
Nyquist diagram/criterion, 324–326
Nyquist-Shannon sampling theorem, 232

Object/target, 3
Offseting circuitry, 345
One-time adjustment, 409
Op-amp’s input (terminal)
  Noninverting/Inverting, 316–318
Open magnetic loop
  Inductive sensors with, 66
Open-loop gain, 252
Operating distance
  Assured, 23
  Effective, 22
  Rated, 22
  Usable, 22
Operating mode
  Active, 225
  Sleep low-power, 225
  Shutdown very low-power, 225
Operating principle, 51
Operation conditions
  Abnormal, 29
  Normal, 29
Operational amplifier (op-amp), 316
  Ideal/Real (commercial), 316–317
  Internally compensated, 326
  Uncompensated, 326
Oscillator
  Bridge, 257, 280–283
  Clapp, 257, 264
  Colpitts, 218, 261–263
  Differential amplifier, 257, 276–280
  Hartley, 257, 260–261

Oscillator *(Continued)*
  Huth-Kuehn, 264
  Meissner, 263
  Pulling-down, 258, 294
  Reinartz, 264
  Relaxation, 157, 257, 306
  Ring/Ring circuit, 213, 257, 265–276
  Three connection-points, 258
Oscillator
  (feedback) LC dipole, 257
  (feedback) LC quadripole, 257, 259, 262
  flat distance characteristic, 258
  loss cancelation, 250
  loss compensation by negative resistance, 257, 285–292, 295–305
  losses, 248
  positive feedback, 257
  with Pulsing DC-current excitation, 257
Oscillatory systems
  Conservative/nonconservative, 242
  Dissipative, 242
  Self-oscilaltory, 242
Output, Antivalent, 13
  Current, 15
  Push-pull (PP) 14, 215
  Ratiometric, 16
  Voltage, 15
Output offset, 342
Output signal range
  Lower limit, 33
  Upper limit, 33
Output stage, Monolithic integrated, 388
Output stage (analog)
  Current, 375
  Ratiometric, 380
  Voltage, 373
Output stage (digital driver), 381–388

Parallel resonant circuit
  LC, LCR, 96, 103–105, 237, 239–242, 290
Passivation, 199
Passive temperature probes, 435
Peak detector, 207, 333–337
Permalloy
  Molybdenum, 182
  Supermalloy, 182
Permalloy (permeable alloy), 137, 181
Permeability (magnetic), 164
  Apparent, 166
  Complex, 167
  Effective, 166
  Imaginary, 167
  Incremental, 167
  Initial, 166
  Relative, 166
  Reversible, 167
Permeability temperature coefficient
  Absolute, 171
  Relative, 172
Permeability temperature factor, 173
Phase margin, 323
Phase path, 245
Phase plane, 245, 246, 250, 286–285
Phase velocity, 245
Phase-locked loop (PLL), 357
  Analog (APLL), 359
  Controllable waveform generator, 358
  Digital (DPLL), 362
  Error/phase detector, 357, 358
  Integer-N-PLL/fractional-N-PLL, 363
  Loop filter, 358
Phase-sensitive detectors
  with analog multiplier, 340–341
  with square-wave reference, 337–340
Phasor diagram, 239
Photolithography (steps)
  Dry etching silicon oxide, 198
  Light exposure, 198
  Photoresist coating, 197
  Photoresist developing, 198
  Reticle aligning, 198
  Wet etching polymerized resist, 198
Pick-and-place, 204
Pilot oscillator, 267
Plane of the states, 245
Polling, 467
Positive feedback (regenerative), 253
Positive temperature coefficient thermistor (PTC), 208, 435, 438
Postprocessing, 119
Power supply
  alternating (AC), 5
  direct current (DC), 5
Precision, 9
Preoperate, 495
Preprocessing, 119
Probe (test) card, 200

Probing
  after bumping, 204
  before bumping, 204
Probing needles, 200
Programmable logic controller (PLC), 474
Protection
  against breaking (open) wire, 29, 376, 400
  against lightning surges, 402–406
  against over-current, 29, 387
  against reverse polarity, 29, 400–402
  against short-circuit, 29, 387
  Thermal overload, 209
  Thermal shutdown, 395, 396
Proximity switches
  capacitive, 20
  inductive, 20
  magnetic, 20
  microwave, 20
  photoelectric, 20
Pt100, Pt1000, 439
PTC activating temperature, 438

Quality factor
  Coil, 79, 102
Quarz crystal (XTAL), 227

Rail-to-rail, 207
Random network access methods:, 468
  Carrier sense multiple access (CSMA), 468
  CSMA with collision avoidance (CSMA/AV), 469
  CSMA with collision detection (CSMA/CD), 469
Range, Bipolar, 6
  Linearity, 6
  Sensing, 6
  Unipolar, 6
  Working, 6
Reactance
  Capacitive, 96, 238
  Inductive, 96, 238
Reading
  Absolute, 5
  Incremental, 5
Real characteristic
  Sample-specific, 8
Rectangular pulse wave, 113
Reference axis, 21
Referred commercial ICs/ASICs
  74HCT9046, CD9046, 363
  AD5291, 5259, 416
  AD532, 341
  AD558, AD7541, 368
  AD590, 443
  AD621-624, 320
  AD694, AD693, 378, 380, 439
  AD7621, ADS7805, 371
  AD849, 266
  ADuC7019, 7029, 476
  AM401, 374
  AM417, 380
  AM462, 376
  AS2701, AS2702, 488, 489
  ASI4U, 489
  BAT42, 266
  DFN6-36, 404
  INA114A, 320
  L6360, L6362A, 496
  LDC1101, 224–231
  LM34, 35, 45, 50, 60, 75, 135, 444
  LM4040, 446
  LM723, 394
  LM741, 394
  LM78xx, LM70xx, 394, 395
  LT1016, 266
  LT1101AC, 320
  NE555, 309–312
  NE566, 308
  PSD36C, 404
  SAP51, SAP5S, 489
  ST232, Max 232, 478
  St485, 479
  TCA505, 212–218
  TDA0161, 218–220
  TDE1708DFT, 218–220
  TIOL111-x, 496
  TL08, 342
  TLV431, SC431, 397, 398
  TPS769xx, 396
  UC27132, 384–386
Reflow oven, 204
Rejustor™
  Functional and auxiliary resistors, 410–411
Reliability
  Intrinsic, 47
  Operational, 47

Reluctance/ magnetic reluctance, 56, 61
Remanent flux density, 165
Repeat accuracy (standard), 35
Repeatabilty, 9
Repeator (network), 466
Report plots, 128
Representative point, 245
Resilience
  Shock, 43
  Vibration, 45
Resistance, 98
  Gyration, 158
  Negative (dynamic), 290, 293, 294
  Switched-off/switched-on, 284
Resistivity (electrical), 164
Resolution, 9
Resolution limit, 10
Resolver
  Sin/cos (vector decomposition), 82
  Standard (electrical), 82
Resonator
  Ceramic, 237
  Mechnical, 242
  Quartz, 237
Responce
  Frequency, 11
  Time, 11
  Transient, 11
Reverse characteristic, 208
Reverse dierection biasing, 208
Reverse transmission, 252, 254
Ripple, Voltage, 34, 392
Rise 7 fall time, 225
RS-232, 477–478
  Data communicatio equipment (DCE), 478
  Data terminal equipment (DTE), 478
RS-422, 477, 478–479
RS-484, 477, 479

Safety margin, 324
Sample deviation, 10
Sample-and-hold (S/H) systems, 341
Sampling interval, 342
Saturation flux density (induction), 164, 165
Saturetion avoidance, 335
Scale factor, 7
Schmitt trigger (regenerative), 207, 308, 355–357
s-domain/s-plane, 254, 255
Section, Frequency
  Determining, 250, 290
  Maintaining, 250, 290
Seeback effect, 51
Segment
  monotonic, 6
Self-teaching adjustments, 417
Sense resistor/voltage, 387
Sensing
  Contact, 4
  Contactless, 4
  Face, 21
  Zone, 21
Sensor, 1
  Embeddable, 21
  flush-mounted, 21
  incremental, 5
  inductive, 2
  linear, 5
  Market, 88
  nonembaddable, 22
  nonflush-mounted, 22
  Sensitivity, 7
Sensor operating principle
  direct energy conversion, 51
  energy flux control, 51
Sensor tests
  routine test, 25
  type test, 25
Sensor trimming
  Functional, 407
  Static/dynamic, 418, 419
  Teach-in, 407, 416, 417, 429
Sensor/actuator industrial networks, 476
Sensus, 1
Serial peripheral interface (SPI), 225
Series resonant circuit
  LC, LCR, 103–104, 237–239
Short-circuit current limitation, 209
Shrinkage, 164
Signal
  Analog, 2
  Binary, 2
  digital encoded, 2
  primary, 2

Signal amplifier, 315
Signal-flow graph, 251
  branch, 251
  Elementary graph reductions, 252
  node, 251
  transmitance, 251
Silicon, 194
Silicon, Doped, 194
Simulation region, 123
Single-pole double-throw switch, 366
Sintering
  Pre-sintering, 175
Sintering (process), 164, 175
Skin effect
  Skin depth, 66, 68, 117
  Skin region, 76
Smith chart, Reactance arcs, 101
  Reactance circles, 101, 102
Snap (independent) action, 28
Software-based compensation with individual self-teaching, 452
Software-defined sensor (SDS), 231
Solenoid, 133
Solid-state device, 190
Solid-state diffusion, 198
Solution data, 126
Solving/ field solving, 119, 126
Source of negative resistance, 285
Span, 6
SPICE simulation, 267, 268, 269, 275, 302–305, 405
Spiral
  Circular, 147
  Hexagonal, 147
  Octagonal, 147
  Square, 147
Square wave
  In-phase, 309
  Quadrature, 309
Standard target, 21, 23
Steinhart-Hart formula, 436
Storage capacitor, 342
Straight line, 7
Straight line Trend, 9
Stray/dispersion flux, 61, 62
Supply voltage
  Dual, 391
  Unipolar, 157, 189, 258, 391
Surface-mount device (SMD), 150
Surface-mount technology (SMT), 150, 203
Susceptance
  Capacitive, 98, 239, 240
  Inductive, 98, 239, 240
Switching function
  Break, 25
  Make, 24
  Make-break (changeover), 25
Synchronous rectifiers, 337
Synchronous serial buses for sensors, 479
  1-wire (SensorPath™ Bus/TI), 479
  2-wire (Inter-Integrated Circuit I2C/ Philips), 479, 489–481
  3-wire (MICROWIRE™ Serial Interface/ NS), 480
  4-wire (Serial Peripheral Interface SPI/ Motorola), 480
Synchronous switching demodulator, 340
System, Stable/unstable, 323
System on chip (SOC), 223

Target, 3, 4
  Cooperative, 4
  Noncooperative, 4
Taylor series representation, 278
Technical core (types), 140, 176–179
Temparature
  Drift, 7
  Storage, 11
Test stimulus, 37
Thermal voltage (VT), 274, 440
Thomsonian circuit, 243
Thomson's formula, 283
Three-terminal fixed voltage regulator, 208
Time
  Checking, 388
  Constant, 11
  Decay, 113
  delay before availability, 27
  Gating, 12
  Lag, 11
  Recovery, 37, 388
  Response, 27
  stability, 7
  Stabilization, 12
Time domain, 254

Tinverse transfer function, 345
Token, 468
Toroid, 135
Transcendetal equation, 336, 441
Transconductance
  Amplifier, 270, 315
  Transistor/Circuit, 276, 290, 315
Transconductance amplifier, 159
Transducer, 1
Transduction
  Mechnism, 2
Transfer admittance, 270
Transfer characteristic, 255
  Dominant pole, 266
  Poles/zeros, 255, 01
Transfer function
Transfer impedance, 270
Transient voltage suppressor diode (TVSD), 209
Transimpedance, 315
  Feedback circuit, 270
Transmission modes, 470
  Asynchronous answer mode (ARM), 471
  Asynchronous balanced mode (ABM), 471
  Normal answer mode (NRM), 470
Transmission rate, 478
Transmitter-receiver (TCVR), 467
Transversal armature, 59
Trench wall processing, 199
Trim factor, 409
Trimming of inductive sensors, 422
  Actuating unit, 422–428, 431–432
  Communication and programming unit, 423, 429–431
  Control unit and peripherals, 422, 428
Trimming potentiometer/trimmer
  Multiturn, 413
  Single-turn, 413
Trimming techniques
  Digital potentiometer, 407, 413–417
  Fusible passive resistor array, 407
  Laser trimmable resistors, 407, 408–410
  Manual mechnical potentiometer, 407–411
Triode valve, 190
Turn off Delay (ToffD), 12, 387
Turn on Delay (TonD), 12, 387
Underwriter's Laboratories
  ISF standard UL913, 47
Unit impulse, 301
Universal Asynchronous Receiver and Transmitter (UART) protocol, 471, 491
Universal serial bus (USB), 16
Uploads adjustment coefficients, 417

Value
  Actual, 9
  Nominal, 9
Van der Pol equation, 306
Variable air gap
  Inductive sensors with, 62
Variable differential transformer
  Linear (LVDT), 72
  Rotary (RVDT), 72
Variable transformation, 347
Variable transformers
  Microsyn, 79, 81
  Resolver, 79, 82
  Synchro, 79, 82
Vector operator (Nabla), 115
Vector products
  Cross product (curl), 115
  Dot product (divergence), 115
Villary effect, 54
Virtual COM port, 229
Voltage
  Rated insulation, 25
  Rated operational, 25
Voltage/current resonance, 238
Voltage amplification, 315
Voltage clamps, 382
Voltage drop, 25
Voltage regulation
  Static/dynamic, 392
Voltage regulators, 391
  Feedback circuit, 392
  Linear, 391
  Low-dropout regulator (LDO), 395
  Pass element (transistor), 392
  Regulating loop, 392
  References (high-accuracy regulators), 391
  Shunt, 391, 398
  Three-terminal adjustable voltage, 395

- Three-terminal fixed voltage, 394, 395

Voltage regulator parameters
- Input regulation factor, 393
- Line regulation, 393
- Load regulation, 393
- Output resistance, 393
- Power supply rejection ratio (PSRR), 393
- Related temperature coefficient (TC), 393
- Temperature coefficient, 393

Voltage-controlled current/voltage source, 159, 207, 265, 271, 315
Voltage-controlled oscillator (VCO), 276
Voltage-current characteristic, 292
Voltage-current characteristic
- letter S shape/letter N shape, 292

Volt-ampere forward characteristic, 208
v-to-i characteristic, 290, 317

Wafer map, 200
Wafer mechanical sawing, 200
Wafer scribing and breaking, 200
Weiss domains, 163
Wheatstone bridge, 345
Winding
- Wire, 134
- Step, 134

Window discriminator, 354
Wire-bonding, 203
Wired output, 3
Wireline/wireless IO-Link, 233
Wireless link, 3

Zener diode, 192, 208
Zero-measurand, 6

# Recent Titles in the Artech House Microelectromechanical Systems (MEMS) Series

*Fundamentals and Applications of Microfluidics, Second Edition*, Nam-Trung Nguyen and Steven T. Wereley

*Hermeticity Testing of MEMS and Microelectronic Packages,* Suzanne Costello, and Marc P. Y. Desmulliez

*Inductive Sensors for Industrial Applications*, Sorin Fericean

*Introduction to Microelectromechanical (MEM) Microwave Systems,* Héctor J. De Los Santos

*An Introduction to Microelectromechanical Systems Engineering*, Nadim Maluf

*MEMS Mechanical Sensors*, Stephen Beeby et al.

*Micro and Nano Energy Harvesting Technologies*, Bin Yang, Huicong Liu, Jingquan Liu, and Chengkuo Lee

*Microfluidics for Biotechnology,* Jean Berthier and Pascal Silberzan

*Organic and Inorganic Nanostructures*, Alexei Nabok

*Post-Processing Techniques for Integrated MEMS,* Sherif Sedky

*RF MEMS Circuit Design for Wireless Communications,* Héctor J. De Los Santos

*Wireless Sensor Network,* Nirupama Bulusu and Sanjay Jha

For further information on these and other Artech House titles, including previously considered out-of-print books now available through our In-Print-Forever® (IPF®) program, contact:

Artech House
685 Canton Street
Norwood, MA 02062
Phone: 781-769-9750
Fax: 781-769-6334
e-mail: artech@artechhouse.com

Artech House
16 Sussex Street
London SW1V 4RW UK
Phone: +44 (0)20 7596-8750
Fax: +44 (0)20 7630-0166
e-mail: artech-uk@artechhouse.com

Find us on the World Wide Web at: www.artechhouse.com